



# Analyse harmonique et Estimation spectrale sur la Sphère.Applications à l'étude du Fond diffus cosmologique.

Frédéric Guilloux

## ► To cite this version:

Frédéric Guilloux. Analyse harmonique et Estimation spectrale sur la Sphère.Applications à l'étude du Fond diffus cosmologique.. Mathématiques [math]. Université Paris-Diderot - Paris VII, 2008. Français. NNT: . tel-00347673

**HAL Id: tel-00347673**

**<https://theses.hal.science/tel-00347673>**

Submitted on 16 Dec 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



U P D – P 7  
U.F.R. M  
E.D. S M P C  
A 2008

T D  
: M A

Présentée par  
**Frédéric GUILLOUX**

---

A E S  
A , F .

---

Thèse dirigée par **Mme Dominique PICARD** et **M. Jean-François CARDOSO**, soutenue publiquement le 8 Décembre 2008.

#### Composition du jury

Rapporteurs :

**M. Andrew JAFFE**, Imperial College London

**M. Jean-Michel LOUBES**, Université Paul Sabatier

Président :

**M. Stéphane MALLAT**, Ecole Polytechnique

Examineurs :

**M. Gérard BIAU**, Université Pierre et Marie Curie

**M. Jean-François CARDOSO**, C.N.R.S.

**M. Jacques DELABROUILLE**, C.N.R.S.

**Mme Dominique PICARD**, Université Paris Diderot – Paris 7

**M. Jean-Luc STARCK**, C.E.A.







# Remerciements

Mes remerciements sont adressés en premier lieu à Dominique Picard et Jean-François Cardoso, pour la confiance qu'ils m'ont accordée tout au long de cette thèse. Ce fut un grand privilège d'apprendre et de travailler sous leur direction. Leur compétence scientifique et leur soutien bienveillant m'ont été précieux pour ces premiers pas dans le monde de la recherche. J'aimerais remercier également Jacques Delabrouille et Gilles Faÿ pour leur encadrement amical et stimulant au sein du laboratoire AstroParticule et Cosmologie, dans lequel j'ai effectué une partie importante de mon travail de thèse.

Andrew Jaffe et Jean-Michel Loubes ont accepté de relire cette thèse et d'en rapporter les résultats. Je leur en suis très reconnaissant. Je remercie également Gérard Biau, Stéphane Mallat et Jean-Luc Starck de l'honneur qu'ils me font en participant à mon jury.

Durant mes études puis mes trois années de monitorat à Chevaleret, j'ai eu la chance de rencontrer des enseignants-chercheurs avec qui j'ai beaucoup apprécié étudier, puis travailler et discuter. En plus de Dominique Picard, qui m'a fait découvrir les statistiques, je remercie particulièrement Michèle Artigue, Olivier Catoni, Emmanuel Germain, Catherine Gille, Gérard Kerkycharian et Erwan Le Pennec. Un grand merci aussi à Laure Elie pour son dévouement à la direction de l'UFR de Mathématiques et à Isabelle, Michèle et Valérie pour leur compétence et leur gentillesse. Mes pensées vont aussi, pour les bons moments passés ensemble et la solidarité, à mes collègues doctorants présents et passés du bureau 5B1 et alentour, spécialement Caroline, Christophe, François, Joseph, Katia, Mohamed, Sophie, Stéphane, Thanh Mai, Thomas, Tu et Vathana.

Ce travail interdisciplinaire a été possible grâce au merveilleux accueil de l'équipe Adamis du laboratoire APC. Je pense notamment à Jacques, Jean-François et Gilles, aux trois benjamins de l'équipe Maude Le Jeune, Héloïse Méheut et Marc Betoule dont l'enthousiasme scientifique est communicatif, ainsi qu'à Dana, Eric, Fabien, Fabrice, Guillaume, Julien, Olivier, Peggy et Radek.

Je remercie également de leur accueil Eric Moulines et les chercheurs de Telecom dont j'ai apprécié la compagnie rue Dareau même si celle-ci fut, de mon fait, trop brève. Merci donc à Céline, François, Gersende, Jamal, Olivier, Yann et Zaïd.

Je souhaite remercier très chaleureusement l'ensemble de l'équipe Modal'X de Nanterre qui m'accueille actuellement et m'a ainsi permis d'achever cette thèse. J'ai trouvé à Nanterre une équipe dynamique et conviviale, avec qui je suis très heureux de travailler. Merci spécialement à Anne-Laure, Annie, Cécile, Christian, Karine, Laurent, Marianne, Mathilde, Olivier, Nathalie, Nathanaël, Patrice, Philippe et Sylvia pour m'avoir aidé à me sentir à l'aise dès le début chez eux. J'en profite pour saluer les doctorants et jeunes docteurs que j'ai été ravi de rencontrer à Nanterre, Ali, Esterina, Hélène, Kaouthar, Marc, Pierre-André et Sébastien, ainsi qu'Olivier que je suis enchanté de retrouver.

J'adresse toute ma reconnaissance à ma famille, à commencer par mes parents qui ont tant fait pour moi et ma soeur Alice. Je pense aussi à mes grands-mères Jeanne et Simone. Je dois beaucoup à mes amis que je ne peux pas tous citer ; qu'ils me le pardonnent. Une pensée tendre pour Amélie.

# Table des matières

<b>Introduction et plan</b>	<b>7</b>
<b>I Présentation du sujet et des résultats</b>	<b>9</b>
<b>1 Le Fond diffus cosmologique</b>	<b>10</b>
1.1 Chronologie de sa découverte et importance dans l’histoire de l’Univers . . . . .	11
1.2 Ses fluctuations : origine et intérêt de leur étude . . . . .	16
1.3 Données disponibles et conséquences pour l’analyse statistique . . . . .	21
<b>2 Analyse sur la sphère</b>	<b>29</b>
2.1 Analyse harmonique dans $\mathbb{L}^2(\mathbb{S}^2, d\xi)$ . . . . .	30
2.2 Champs aléatoires sur la sphère . . . . .	35
2.3 Ondelettes sphériques . . . . .	39
<b>3 Enjeux statistiques liés à l’étude du Fond diffus cosmologique</b>	<b>47</b>
3.1 Modélisation du Fond diffus et des observations . . . . .	48
3.2 Estimation du spectre de puissance angulaire . . . . .	50
3.3 Extraction du Fond diffus (séparation de source) . . . . .	55
3.4 Autres enjeux statistiques . . . . .	58
<b>4 Présentation des résultats obtenus</b>	<b>61</b>
4.1 Optimisation non-asymptotique des needlets (chapitre 5) . . . . .	62
4.2 Nouvelle méthode d’estimation spectrale (chapitres 6 et 7) . . . . .	66
4.3 Applications au Fond diffus cosmologique (chapitres 6 et 8) . . . . .	73
4.4 Note technique . . . . .	77
4.5 Perspectives . . . . .	77

<b>II</b>	<b>Articles de revues</b>	<b>79</b>
<b>5</b>	<b>Practical wavelet design on the sphere</b>	<b>80</b>
	Introduction . . . . .	81
5.1	Needlets frames . . . . .	82
5.2	Design of optimally localized wavelets . . . . .	87
5.3	Examples, numerical results . . . . .	96
5.4	Conclusion . . . . .	100
5.5	Proofs . . . . .	101
<b>6</b>	<b>CMB power spectrum estimation using wavelets</b>	<b>103</b>
6.1	Framework . . . . .	106
6.2	The needlet spectral estimators (NSE) . . . . .	109
6.3	Monte Carlo studies . . . . .	114
6.4	Discussion . . . . .	125
6.5	Conclusion . . . . .	126
6.6	Appendices . . . . .	126
<b>7</b>	<b>Spectral estimation on the sphere with needlets : high frequency asymptotics</b>	<b>130</b>
7.1	Introduction . . . . .	131
7.2	Model and settings . . . . .	132
7.3	Estimation results . . . . .	136
7.4	Simulations . . . . .	139
7.5	Conclusion . . . . .	142
7.6	Proofs . . . . .	143
<b>8</b>	<b>A full sky, low foreground, high resolution CMB map from WMAP</b>	<b>152</b>
8.1	Introduction . . . . .	153
8.2	General considerations . . . . .	155
8.3	Evaluation and comparison of available maps . . . . .	160
8.4	The ILC needlet method . . . . .	168
8.5	Simulations . . . . .	177
8.6	Application to WMAP data . . . . .	179
8.7	Discussion . . . . .	181
8.8	Conclusions . . . . .	191
8.9	Derivation of the ILC bias . . . . .	192
	<b>Bibliographie</b>	<b>200</b>

# Introduction et plan

Cette thèse est le fruit d'un travail interdisciplinaire, au croisement des statistiques, de l'analyse harmonique et de l'analyse de données en cosmologie. Elle contribue à l'étude statistique de champs aléatoires sphériques. L'objectif applicatif est d'apporter des améliorations dans l'analyse des données d'observations du Fond diffus cosmologique. Pour cela, des résultats récents d'analyse sur la sphère sont utilisés et enrichis.

L'étude du Fond diffus cosmologique est un des piliers de la cosmologie moderne. Son analyse soulève des questions statistiques attrayantes et originales, à commencer par la géométrie (sphérique) des observations et la complexité des modèles mis en jeu. Le traitement des données du Fond diffus est traditionnellement mené en plusieurs étapes, qui font chacune intervenir des problèmes statistiques particuliers. Les principales (mais non les seules) sont 1) la fabrication de cartes sphériques à partir des séries temporelles des observations ; 2a) la séparation des différentes sources superposées (dont le Fond diffus) ; 2b) l'estimation du spectre de puissance angulaire ; 3) l'estimation des paramètres du modèle cosmologique sous-jacent. Il sera principalement question ici des étapes 2a) et 2b), c'est-à-dire celles impliquant l'analyse de champs aléatoires sur la sphère.

Du point de vue statistique, la démarche a été de travailler en partant d'une situation physique réelle et en se confrontant en permanence aux données, notamment dans le choix des modèles et des hypothèses. Cette approche a aussi nécessité la mise en oeuvre numérique des outils mathématiques et procédures statistiques proposées, d'une façon compatible avec les vastes jeux de données complexes auxquelles on s'intéresse.

Bien que le travail dans son ensemble soit de nature principalement statistique, plusieurs chapitres présentent des éléments d'analyse et de physique nécessaires pour comprendre les méthodes employées et les enjeux.

Dans le chapitre 1, l'étude du Fond diffus cosmologique, et en particulier de son spectre de puissance, est motivée physiquement et replacée dans son contexte historique. L'accent est mis aussi sur les conditions de son observation en vue de leur modélisation statistique.

Le chapitre 2 fournit les concepts et outils de bases de l'analyse harmonique sur la sphère pour des fonctions déterministes et aléatoires. Il se conclut par la construction de frames d'ondelettes sur la sphère, dont l'utilisation est l'une des originalités de cette étude.

Le chapitre 3 utilise les deux précédents pour présenter certains problèmes statistiques fondés sur l'observation de champs aléatoires sphériques et directement motivés par l'application au Fond diffus. Le principal est l'estimation du spectre de puissance.

Le chapitre 4 résume les résultats qui sont détaillés ensuite, et propose des pistes pour les prolonger.

Les quatre derniers chapitres sont des articles de revues (dont l'un est paru, deux sont acceptés et le dernier est soumis). Ils sont écrits en collaboration avec des membres de l'équipe ADAMIS<sup>1</sup>, au sein de laquelle j'ai réalisé l'essentiel de ma recherche au cours de ma thèse.

Le chapitre 5 présente une étude sur les needlets et leur optimisation. Les needlets sont des fonctions sphériques à symétrie axiale dont les "rotationnées" et dilatées (la dilatation étant définie dans l'espace de Fourier) constituent une frame. Nous considérons des needlets à bande limitée (dans l'espace de Fourier) et optimisons leur localisation spatiale au sens de deux critères. Le premier est la concentration de leur énergie (de leur norme) et le second est leur insensibilité à des données manquantes dans l'analyse d'un champ stationnaire. Dans les deux cas, nous montrons que la quantité à minimiser est une forme quadratique des filtres utilisés. Nous complétons cette étude par son implémentation et par des comparaisons numériques avec d'autres approches.

Le chapitre 6 introduit une nouvelle méthode d'estimation du spectre de puissance d'un champ sphérique stationnaire, en l'occurrence le Fond diffus cosmologique. L'estimation est réalisée à partir des coefficients de needlet des observations, ce qui permet aux estimateurs de s'adapter au niveau local du rapport signal/bruit et de traiter le cas d'un modèle hétéroscédastique avec des données manquantes. Deux séries de simulations en illustrent les performances : la première pour un modèle à une seule observation et la seconde pour des données longitudinales fournies par plusieurs observations. Dans le premier cas, les estimateurs proposés améliorent les techniques existantes et réalisent automatiquement une transition entre différents régimes traités habituellement de façon ad hoc. La capacité à traiter le second cas (données longitudinales), est une spécificité de notre méthode. Le modèle réaliste que nous utilisons comporte des données manquantes se recoupant partiellement entre observations, et des niveaux de bruit (inhomogènes), résolutions angulaires et opérateurs de convolution tous différents d'une observation à l'autre.

Le chapitre 7 reprend les mêmes estimateurs spectraux par needlet, mais d'un point de vue différent : nous nous intéressons à leurs propriétés asymptotique au sens des hautes fréquences, plus précisément à leur convergence, qui prend dans ce contexte un sens particulier (puisque l'on ne considère qu'une réalisation *unique* du champ aléatoire et une suite de paramètres différents). Nous obtenons deux résultats de convergence, dont l'un, à condition de connaître l'ordre de grandeur du spectre, autorise un niveau de bruit arbitrairement grand dans certaines régions.

Le chapitre 8 est une autre application des needlets, dans un contexte statistique différent mais lui aussi d'une grande importance pour l'étude du Fond diffus cosmologique. Il s'agit de la séparation de sources : étant données plusieurs observations bruitées de mélanges d'un champ stationnaire (le Fond diffus) et d'autres sources (champs aléatoires modélisant les autres composantes astrophysiques) supposées indépendantes, l'objectif est d'extraire le Fond diffus des observations. Après une étude sur des simulations et une comparaison détaillée avec les principales autres méthodes existantes, notre méthode est appliquée aux données de l'expérience WMAP<sup>2</sup>.

---

<sup>1</sup>Analyse de Données en Astroparticule, Modélisation, Interprétation, Simulation ; équipe du laboratoire APC (Astro-Particule et Cosmologie)

<sup>2</sup>Wilkinson Microwave Anisotropy Probe : la principale expérience actuelle dédiée à l'observation du Fond diffus cosmologique.

## **Première partie**

# **Présentation du sujet et des résultats**

# Chapitre 1

## Le Fond diffus cosmologique

### Sommaire

1.1	Chronologie de sa découverte et importance dans l'histoire de l'Univers . . . . .	<b>11</b>
	Du modèle standard . . . . .	11
	... au Fond diffus . . . . .	13
	Origine physique du rayonnement . . . . .	14
1.2	Ses fluctuations : origine et intérêt de leur étude . . . . .	<b>16</b>
	Fluctuations primaires . . . . .	16
	Fluctuations secondaires . . . . .	18
	Avant-plans . . . . .	19
1.3	Données disponibles et conséquences pour l'analyse statistique . . . . .	<b>21</b>
	Exemples d'expériences passées et en cours . . . . .	23
	Exemples d'expériences programmées . . . . .	24
	Conséquences pour l'analyse statistique . . . . .	26

La cosmologie est l'étude de l'Univers dans son ensemble, notamment son organisation et son évolution... depuis que l'on sait qu'il en a une. L'un des bouleversements scientifiques majeurs du vingtième siècle fut en effet la découverte que l'Univers a une histoire, durant laquelle les formes d'existence de la matière ont été variées.

Certaines phases de ce mouvement sont maintenant connues, grâce d'une part à des observations célestes et des expérimentations physiques de plus en plus précises, et d'autre part aux théories issues d'autres domaines de la physique et des mathématiques. Il est ainsi possible de retracer, dans les grandes lignes et parfois même avec précision, certaines étapes de l'évolution cosmique datant de plusieurs milliards d'années.

Un élément fondamental dans cette compréhension de l'Univers est le Fond diffus cosmologique, dont l'acronyme anglais est *CMB* (pour Cosmic Microwave Background). Sa découverte en 1965 a validé la théorie d'un Univers en expansion à partir d'un état extrêmement chaud et condensé, qui est aujourd'hui le *modèle standard* de la cosmologie. Depuis cette découverte, l'observation du CMB est l'objet d'un effort considérable de la communauté scientifique, comme en témoignent les nombreuses expériences passées et programmées dont il sera question plus loin.

Les données issues de ces observations sont maintenant tellement nombreuses et précises qu'elles permettent, au-delà de la validation d'un modèle théorique général, d'estimer les paramètres de ce

modèle, avec une incertitude descendant parfois jusqu’à quelques pour cents, au point que les cosmologistes n’hésitent pas à parler de l’entrée dans l’ère de la *cosmologie de précision*.

Dans le même temps, la quantité, la variété et la complexité de ces données observationnelles nécessitent une analyse, en particulier une analyse statistique, de plus en plus élaborée ou, si l’on veut, “de précision” elle aussi. Un témoin de cette nécessité est l’intérêt croissant, ces toutes dernières années, de scientifiques issus des statistiques mathématiques et du traitement du signal, pour l’étude des données astrophysiques en général et celles liées à l’observation du Fond diffus cosmologique en particulier.

Dans ce premier chapitre consacré au *CMB*, tentons, sans entrer dans le détail des théories physiques mises en jeu, de faire ressortir la place du Fond diffus dans la cosmologie moderne ainsi que les enjeux et difficultés de son étude statistique (dont le formalisme mathématique est reporté au chapitre 3). Pour des exposés récents consacrés au Fond diffus cosmologique incluant une présentation des équations physiques, on pourra consulter, par exemple, les articles de Hu et Dodelson [90] ou Kozowsky [97], le livre de Dodelson [72] ou en français le cours de Lachièze-Rey [98].

## 1.1 Chronologie de sa découverte et importance dans l’histoire de l’Univers

### Du modèle standard ...

Le modèle standard, qui s’est développé tout au long du vingtième siècle, est né dans les années 1920 d’une observation et d’un cadre théorique : l’observation de la récession des galaxies et la théorie de la relativité. A cette époque en effet les astronomes établissent, grâce à des télescopes toujours plus puissants, que les “nébuleuses” qu’ils observent depuis longtemps sont en réalité des galaxies semblables à la nôtre, sujet discuté depuis Kant. De plus, Hubble mesure le spectre lumineux de ces galaxies et montre que les longueurs d’onde qui nous parviennent sont d’autant plus grandes que les galaxies sont lointaines. Si l’on admet que la matière (en l’occurrence les atomes dont on mesure les raies d’absorption et d’émission) a partout les mêmes propriétés physiques, alors ce *décalage vers le rouge* des spectres lumineux des galaxies peut être interprété comme un effet Doppler. Partant de ceci, Hubble met en évidence une tendance linéaire entre la vitesse d’éloignement des galaxies et la distance qui les sépare de nous (le rapport de proportionnalité est depuis appelé *constante de Hubble* et noté  $H_0$ ). Si l’on admet le *principe cosmologique* selon lequel il n’y a pas de lieu ni de direction privilégiés dans l’Univers<sup>1</sup>, cette observation suggère un mouvement général d’expansion.

Si cette découverte est à l’origine d’une transformation de notre vision du monde, c’est parce qu’elle a pu s’inscrire dans un modèle théorique permettant la description du mouvement à l’échelle de l’Univers. Ce cadre théorique, la relativité, est mis au point par Einstein quelques années auparavant. Il permet à Friedmann, puis à Lemaître, dans les années 1920, d’établir des équations décrivant un Uni-

---

<sup>1</sup>Le principe cosmologique est le postulat sur lequel repose toute la cosmologie moderne. Il stipule l’Univers homogène (ses caractéristiques physiques sont les mêmes en tout point de l’espace) et isotrope (son aspect ne dépend pas de la direction dans laquelle on l’observe). Ce principe cosmologique est à comprendre aux grandes échelles. Aux échelles du système solaire ou de la Voie Lactée il n’est clairement pas vérifié. Par contre, les grands sondages de galaxies montrent depuis une vingtaine d’années qu’aux échelles les plus grandes qui nous sont accessibles aujourd’hui le principe cosmologique est validé : au-delà de quelques milliards d’années-lumière les structures s’estompent. Par ailleurs, l’homogénéité du Fond diffus cosmologique permet elle aussi de valider a posteriori le principe cosmologique.



vers en évolution, en fonction de son contenu. Plus explicitement, et à titre d'illustration uniquement, introduisons le facteur d'échelle  $a$  qui est une fonction du temps intervenant à chaque instant comme un coefficient multiplicatif des coordonnées d'espace. Une fonction  $a(t)$  croissante correspond à un Univers en expansion, décroissante à un Univers en contraction et constante à un Univers statique. Ainsi la constante de Hubble n'est autre que  $H_0 = \frac{a'}{a}(t = \text{aujourd'hui})$ . Alors, en appliquant la théorie d'Einstein à un fluide homogène, on obtient les *équations de Friedmann* :  $\left(\frac{a'(t)}{a(t)}\right)^2 = -\frac{K}{a^2(t)} + \frac{8\pi G}{3}\rho(t) + \frac{\Lambda}{3}$  et  $\frac{a''(t)}{a(t)} = -\frac{4\pi G}{3}(\rho(t) + 3p(t)) + \frac{\Lambda}{3}$ , où  $p$  et  $\rho$  sont respectivement la pression et la densité du fluide,  $G$  est la constante de Newton,  $K$  est la courbure spatiale de l'espace à 3 dimensions et  $\Lambda$  est une constante, la "constante cosmologique", dont le statut est encore en discussion. Une caractéristique de ces équations est que, pour un large domaine de leurs paramètres, leurs solutions  $a$  sont des fonctions passant par 0 en un temps passé fini.

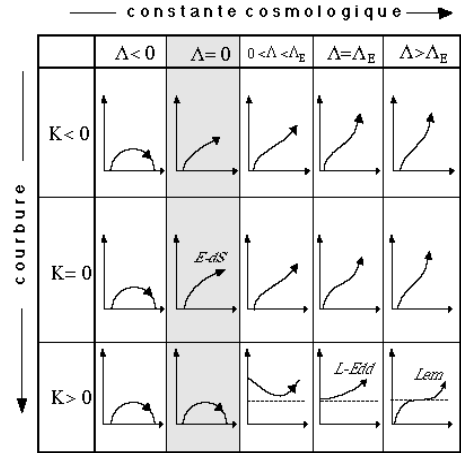
Ces équations reposent sur des simplifications et idéalisations, comme le fait que l'Univers à grande échelle soit modélisé par un fluide homogène. Et surtout, la limite d'une singularité initiale suppose que l'extrapolation des notions et théories physiques connues est possible jusqu'à cet "instant". Néanmoins, leur efficacité à décrire, au moins approximativement, les observations cosmiques lointaines et à les relier à des phénomènes physiques connus ou du moins étudiables, en ont fait la base de tout le développement ultérieur du modèle standard. Aujourd'hui, les physiciens invoquent trois piliers observationnels soutenant ce modèle d'Univers en expansion. Tout d'abord, le décalage vers le rouge de la lumière des galaxies, dont le modèle peut rendre compte et qui, depuis Hubble, est confirmé pour des galaxies toujours plus lointaines. En second lieu, les abondances relatives des différents éléments chimiques observés dans l'Univers, interprétées comme résultant d'une synthèse primordiale des éléments légers avant l'apparition des étoiles et galaxies, puis d'une synthèse des éléments plus lourds dans les étoiles (et les explosions d'étoiles ou supernovae). En effet, la formation d'éléments telle qu'elle se produit aujourd'hui est incompatible avec les abondances observées si l'on ne suppose pas une telle synthèse primordiale à une époque plus ancienne de l'histoire de l'Univers. Ce point ne sera plus abordé dans la suite de ce travail. Le troisième pilier du modèle standard est le Fond diffus cosmologique, dont ce modèle a permis la prédiction et dont ni les propriétés ni même l'existence ne trouvent d'explication satisfaisante en dehors de ce modèle.

Jusqu'en 1965 et l'annonce de l'observation du *CMB*, ce qui deviendra le modèle "standard" de la cosmologie est encore controversé. Certains scientifiques (dont Einstein dans un premier temps) préfèrent à ce modèle de "gros boum"<sup>2</sup> celui d'un univers stationnaire sans évolution d'ensemble, c'est-à-dire sans histoire. Mais les modèles d'univers stationnaires proposés nécessitent, pour rendre compte des observations, des hypothèses problématiques comme la création continue de matière. Surtout, ces modèles ne permettent pas d'expliquer le Fond diffus cosmologique (ni l'abondance d'éléments légers). L'observation en 1965 du *CMB* qui, à l'inverse, avait été prédit par la théorie de l'Univers en expansion, a rallié la majorité des physiciens à cette théorie.

---

<sup>2</sup>Le terme *Big Bang* fut introduit pour qualifier par dérision le modèle d'un Univers en expansion, par un adversaire de ce modèle (Fred Hoyle)... avant de connaître le succès que l'on sait. Ce terme contribue peut-être à des interprétations pour le moins abusives de la théorie, en termes d'"instant initial", d'"explosion", de "création", etc.

Terminons cette brève présentation du modèle standard en précisant que les équations de ce dernier (celles de Friedmann et d'autres) dépendent d'un certain nombre de paramètres, appelés *paramètres cosmologiques*. La figure 1.1 illustre à quel point les solutions des équations de Friedmann sont sensibles à ces paramètres. La détermination de ces paramètres est donc cruciale pour la description de l'évolution de l'Univers. Une fois le modèle admis, c'est même, pour ainsi dire, le but cosmologique ultime de toute expérience et de toute analyse. Bien que le nombre de paramètres cosmologiques dépende des variantes du modèle standard, ce nombre reste généralement de l'ordre de la dizaine. Parmi ceux-ci, intervenant clairement comme paramètres ou conditions initiales des équations de Friedmann ci-dessus, citons entre autres : la constante de Hubble  $H_0$  ; la densité moyenne de matière aujourd'hui  $\rho_0$ , souvent exprimée à travers la quantité  $\Omega_m := \frac{8\pi G\rho_0}{3H_0^2}$  ; la courbure  $K$  ou bien  $\Omega_K := \frac{-K}{H_0^2 a^2(0)}$  ; la constante cosmologique  $\Lambda$  (ou, sous sa forme normalisée,  $\Omega_\Lambda := \frac{\Lambda}{3H_0^2}$ ) dont les physiciens donnent parfois une interprétation physique en terme d'“énergie noire”.



F . 1.1 – Solutions  $t \mapsto a(t)$  des équations de Friedmann selon la valeur de deux paramètres cosmologiques.

### ... au Fond diffus

Comme on le voit, le modèle standard, bien que partant d'hypothèses relativement simples, est un modèle riche. Les physiciens ont tenté de tirer de ce modèle le maximum d'implications, même du temps où il était controversé. C'est ainsi que dans les années 1940, l'étude ne se limite plus à une description géométrique de l'évolution de l'Univers. On étudie aussi la physique de l'Univers primordial. Le nom de Gamow est associé à la description des propriétés physiques que devait avoir l'Univers à une époque où le facteur d'échelle  $a(t)$  était très petit. Si l'Univers peut être modélisé par un gaz parfait, alors quand sa densité était plus grande sa température devait l'être aussi. Puis, comme un gaz qui se détend, il s'est refroidi avec l'expansion.

Ce raisonnement a de nombreuses conséquences. L'une d'elles, détaillée dans le prochain paragraphe, est l'émission, au moment où la température de l'Univers passe en dessous d'une certaine température critique, d'un rayonnement qui, se propageant dans tout l'Univers, doit être visible encore aujourd'hui. En effet, les photons émis à cette époque l'ont été de tous les endroits de l'Univers. Donc, ceux qui ont été émis à une distance de nous suffisante, après avoir voyagé depuis ce temps dans l'espace, doivent nous parvenir aujourd'hui. Cette prédiction théorique est établie en 1948 par Alpher, Hermann et Gamow. Ils démontrent également que le rayonnement émis doit avoir un spectre lumineux de *corps noir*, c'est-à-dire que la distribution de sa luminance énergétique  $B$  selon les longueurs d'onde  $\lambda$  suit la loi de Planck :  $B(\lambda, T_0) = \frac{2hc^2}{\lambda^5} \left( \exp\left(\frac{hc}{kT_0\lambda}\right) - 1 \right)^{-1}$ , où  $c$  est la vitesse de la lumière,  $h$  et  $k$  sont respectivement les constantes de Planck et de Boltzmann. Cette distribution selon  $\lambda$  dépend du paramètre  $T_0$ , la température du corps émettant le rayonnement. Il est facile de constater (voir l'annexe en fin de chapitre) que lorsqu'on tient compte de l'expansion de l'Univers pendant le parcours d'un tel rayonnement, le spectre lumineux observé est également un spectre de corps noir, avec une température apparente  $T(t) = \frac{a(0)}{a(t)} T_0$ . C'est pourquoi, quelles que soient les méthodes de mesures, les

physiciens décrivent tout rayonnement par sa température<sup>3</sup>.

Les auteurs précités calculent que la température apparente du rayonnement cosmologique doit être de quelques degrés Kelvin, donc l'intensité de ce rayonnement est faible (comparée aux sources lumineuses qui nous entourent), et maximale pour des longueurs d'onde supérieures à celles de la lumière visible (plus précisément dans le domaine des micro-ondes, d'où la dénomination anglophone). D'après le principe cosmologique, les propriétés de ce rayonnement, c'est-à-dire sa température, devraient être les mêmes quelle que soit la direction d'observation.

Dans ces mêmes années, la radio-astronomie ("observant" le ciel dans des longueurs d'onde plus grandes que le visible) prend son essor. Le rayonnement prédit par la théorie est observé plusieurs fois par des astronomes qui n'en perçoivent pas la signification. Ainsi, sur la carte du ciel (reproduite figure 1.15) présentée en 1955 par Denisse, Le Roux et Steinberg [70] relatant leurs mesures dans le domaine des ondes radio, on voit clairement la surface de niveau à 3 Kelvin en dehors de la Voie Lactée. Mais ce n'est qu'*a posteriori* qu'on peut l'interpréter comme une détection précoce du Fond diffus cosmologique. La découverte<sup>4</sup> de ce dernier revient à Penzias et Wilson (pour l'observation) et à Dicke, Peebles, Roll et Wilkinson (pour l'interprétation) dans deux articles publiés conjointement en 1965 [120, 71]. Penzias et Wilson mesuraient le rayonnement céleste à la longueur d'onde de 7 cm avec une antenne originellement dédiée aux communications satellitaires. Ils reportent un bruit électromagnétique constant dans le temps et selon les directions, avec une puissance, dans la longueur d'onde observée, correspondant à celle d'un corps noir à  $3.5 \pm 1.0$  K. Dicke et al. interprètent ce signal comme le Fond diffus cosmologique, dans le cadre d'un modèle sans singularité initiale mais où l'Univers a connu dans le passé un état suffisamment chaud et dense.

## Origine physique du rayonnement

Reprenons le raisonnement du paragraphe précédent sur l'histoire thermique de l'Univers pour expliquer le processus qui a produit ce rayonnement. Si l'on remonte suffisamment loin dans le passé, il doit exister une époque où l'Univers est suffisamment chaud pour que l'agitation thermique interdise l'existence d'atomes stables. La matière est donc ionisée : les électrons sont libres, non rattachés aux noyaux. Les physiciens parlent d'un *plasma*, au sein duquel les photons interagissent fortement avec les électrons et les noyaux d'hydrogène et d'hélium (sans parler des autres particules). Les photons sont absorbés et réémis continuellement par les électrons et les noyaux, avec un libre parcours moyen très faible, de sorte que toutes ces particules sont à l'équilibre thermique et évoluent dynamiquement ensemble. Cet équilibre est possible car la durée moyenne entre deux interactions photons – électrons ou photons – noyaux est courte devant l'échelle de temps de l'expansion, qui est aussi celle du refroidissement. En particulier, la matière ne peut pas se condenser sans entraîner le rayonnement. Or celui-ci possède une grande capacité de résistance à la compression. Sa présence empêche donc la formation d'étoiles ou de galaxies.

Lorsque la température baisse suffisamment (un peu plus de 3000 K), deux phénomènes modifient profondément et relativement brutalement l'organisation de la matière : la recombinaison et le découplage. La *recombinaison*<sup>5</sup> désigne la fin de l'ionisation et la formation d'atomes neutres par l'association des électrons avec les noyaux, association que l'agitation thermique est maintenant insuffisante à

<sup>3</sup>Même quand un rayonnement n'a pas le spectre lumineux d'un corps noir, la donnée fournie est la température équivalente que devrait avoir un corps noir pour émettre avec la même intensité à la longueur d'onde mesurée

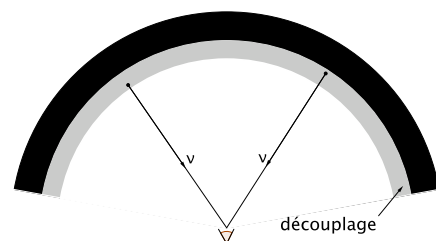
<sup>4</sup>Pour cette découverte, Penzias et Wilson ont obtenu le prix Nobel de physique en 1978.

<sup>5</sup>Selon notre modèle actuel, cette combinaison a lieu à ce moment pour la première fois. Le préfixe "re-" du terme,

casser. Alors, du fait de la chute de la densité des électrons libres et des noyaux ionisés, les photons n’interagissent plus avec le reste de la matière et se propagent librement dans l’espace. C’est le *dé-couplage*. On dit que l’Univers devient transparent à son propre rayonnement, tandis que le plasma était auparavant opaque. Après le découplage, les atomes et le rayonnement vont évoluer (en première approximation) indépendamment. Les premiers, soumis à la gravitation, vont former les grandes structures (amas, galaxies) que l’on voit plus près de nous. Le rayonnement, au contraire, pourra désormais se propager librement à travers l’espace. Ainsi, le Fond diffus cosmologique est le rayonnement électromagnétique le plus ancien (et donc aussi le plus lointain) qui pourra jamais nous arriver. En effet, nul rayonnement plus ancien n’a pu se propager librement dans l’Univers avant le découplage. C’est donc la limite de l’Univers “visible”, ainsi qu’une trace de l’Univers antérieure aux premières étoiles et galaxies.

Ce mécanisme explique que le rayonnement émis soit celui d’un corps noir, il est en effet précisément celui de photons en équilibre thermique avec la matière. Les observations ont depuis confirmé que le spectre lumineux de ce rayonnement est conforme à la loi de Planck. C’est même le spectre de corps noir le plus parfait jamais mesuré<sup>6</sup>. Sa température apparente est de 2.726 K (ce qui, compte tenu de la température d’émission, indique un taux d’expansion  $\frac{a(t)}{a(0)} \simeq 1100$  depuis cette époque) et elle est très homogène sur le ciel. Au niveau de précision des mesures des années 1960-70, elle pouvait même sembler parfaitement constante dans toutes les directions.

Si la recombinaison et le découplage s’étaient produits instantanément, la zone d’émission des photons nous arrivant aujourd’hui serait une sphère (de quelques 14 milliards d’années-lumière de rayon), appelée *surface de dernière diffusion*. Les physiciens estiment qu’en réalité la durée de ces phénomènes a été d’environ 100 000 ans. Bien que les données d’observations soient sphériques (puisqu’indexées par des directions d’observations), les valeurs mesurées correspondent donc à des moyennes sur la petite épaisseur (comparée au rayon) de la surface de dernière diffusion.



F . 1.2 – Surface de dernière diffusion.

A partir de sa découverte en 1965, l’observation et l’analyse du fond diffus cosmologique ont constitué un domaine de recherche à part entière. Les premières études ont visé à vérifier le spectre lumineux de corps noir, ce qui fut fait avec une extrême précision. Puis, on chercha à détecter des inhomogénéités spatiales dans le rayonnement, c’est-à-dire des fluctuations de sa température selon la direction observée, ce dont il est question dans la section suivante.

Terminons cette section par une remarque méthodologique. Le travail présenté ici s’inscrit dans le cadre du modèle standard de la cosmologie, qui est celui adopté par les physiciens, sans prétendre porter un jugement ou chercher à contribuer à sa validation. La présentation donnée dans cette section [1.1] et dans la suivante [1.2] du modèle standard et l’interprétation qu’elle permet du rayonnement de fond observé, sont utiles pour motiver l’étude qui suit et la replacer dans un contexte scientifique bien plus large. Notons cependant que les méthodes mathématiques d’analyse et d’estimation sur la sphère dont il sera question ensuite sont largement indépendantes de ces modèles physiques (seules sont parfois supposées l’isotropie et la gaussianité des fluctuations, définies dans la suite).

quoique consacré par l’usage, est donc superflu.

<sup>6</sup>Jonh Mather a obtenu le prix Nobel de physique en 2006 pour la mesure de ce spectre grâce à l’expérience COBE.

## 1.2 Ses fluctuations : origine et intérêt de leur étude

L'étude actuelle du Fond diffus cosmologique est celle de ses inhomogénéités spatiales, appelées *anisotropies* par les physiciens. Ces fluctuations des propriétés du rayonnement (sa température, mais aussi sa polarisation) en fonction de la direction d'observation sont assez faibles pour avoir été négligées dans un premier temps. L'importance capitale néanmoins de ces fluctuations pour notre compréhension de l'histoire cosmique est détaillée dans la suite. Mentionnons simplement dès maintenant que de telles inhomogénéités sont nécessaires pour expliquer la formation des étoiles, galaxies et amas de galaxies. En effet, comment expliquer l'apparition de ces structures, qui représentent une organisation spatiale si inhomogène de la matière, si le milieu dont elles sont issues était parfaitement homogène ? Il n'y a pas encore de scénario solidement établi expliquant l'origine de fluctuations dans l'Univers primordial. La théorie dite de "l'inflation" fournit celui qui semble actuellement le plus largement partagé par les physiciens. Elle prédit de petites fluctuations aléatoires de densité suivant une loi gaussienne stationnaire. Une fois admise l'existence de fluctuations initiales, les lois régissant leur évolution dans le plasma avant découplage sont en revanche bien connues. Au moment du découplage, l'évolution des fluctuations est gelée, si bien que le Fond diffus cosmologique doit garder la trace de l'état des fluctuations à cet instant.

Peu de temps après l'observation du rayonnement lui-même, Silk [131] calcula l'effet d'hypothétiques petites fluctuations de densité dans le plasma (avant le découplage) sur l'observation du Fond diffus. Ces calculs, qui supposaient un découplage instantané, furent ensuite généralisés en tenant compte de l'épaisseur de la surface de dernière diffusion et de nombreux effets indirects. Par exemple, Sachs et Wolfe [127] ont étudié l'effet d'inhomogénéités de densité le long du trajet du rayonnement après son émission ; Sunyaev et Zeldovich [136] ont calculé l'effet des interactions entre le rayonnement et des électrons libres sur ce même trajet. La première observation<sup>7</sup> fiable des fluctuations du Fond diffus a été réalisée en 1992 par le satellite COBE [132]. L'ordre de grandeur relatif des fluctuations de la température du rayonnement selon la direction observée est de  $10^{-5}$  par rapport à la valeur moyenne (hormis une fluctuation dipolaire plus importante).

A l'inverse de la section précédente, commençons par indiquer les principaux phénomènes physiques à l'origine de ces fluctuations. La prochaine section [1.3] donnera des exemples d'expériences dédiées à leur observation. On distingue les fluctuations primaires, correspondant à des inhomogénéités au moment de l'émission sur la surface de dernière diffusion, et les fluctuations secondaires provoquées par des phénomènes physiques postérieurs à la recombinaison. Les premières nous renseignent donc sur l'Univers d'avant la recombinaison, tandis que les secondes apportent des informations sur l'Univers plus récent. Enfin, des avant-plans, situés entre nous et la surface de dernière diffusion, émettent des rayonnements qui se superposent aux fluctuations du Fond diffus.

### Fluctuations primaires

Des inhomogénéités initiales du plasma évoluent sous les actions opposées de la gravité et de la pression de radiation des photons. La gravité tend à accentuer les inhomogénéités, à faire s'effondrer la matière vers les régions les plus denses. Au contraire, la pression de radiation des photons tend à uniformiser la matière en détendant les régions les plus denses. La compétition entre ces deux forces aboutit à des oscillations acoustiques qui amplifient des perturbations de densité initiales même

---

<sup>7</sup>Pour cette observation, Smoot a obtenu le prix Nobel de physique en 2006.

infimes. La période des oscillations dépend de la taille des fluctuations et des propriétés physiques du milieu et de l'espace.

Ce processus est bloqué au moment du découplage et le rayonnement émis possède une température qui varie selon la région d'émission. Ces variations (faibles, rappelons-le, par rapport à la valeur moyenne, car le plasma est resté malgré tout très homogène) sont la résultante de plusieurs effets :

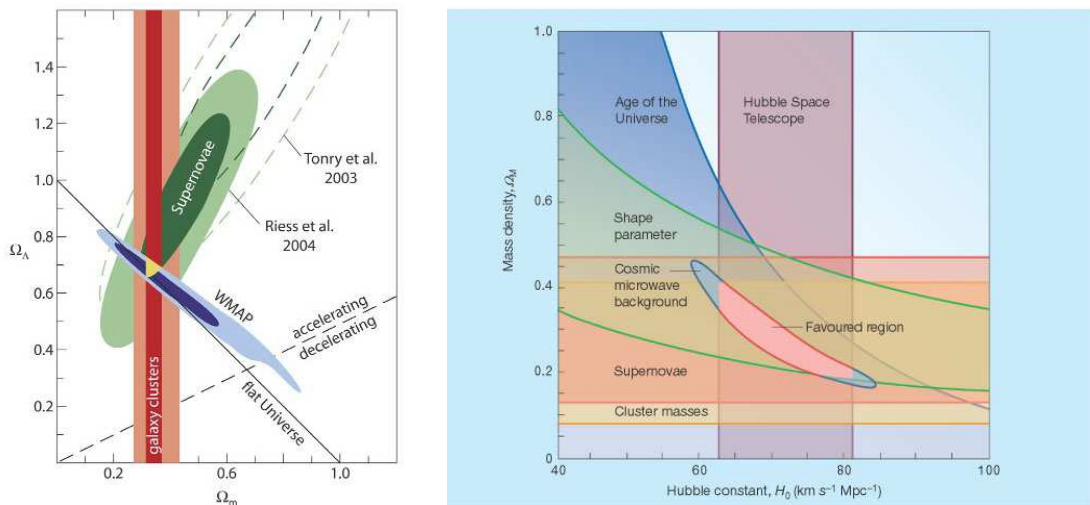
- La température intrinsèque du milieu d'émission dépend de sa densité. Comme pour un gaz, les régions plus denses sont plus chaudes.
- Cet effet est contrebalancé par un effet relativiste, dit effet Sachs-Wolfe : la longueur d'onde des photons est modifiée par la densité plus ou moins grande de la région qu'ils quittent. Les régions plus denses font perdre plus d'énergie aux photons qui les quittent, de sorte que leur longueur d'onde augmente (autrement dit, leur température équivalente diminue).
- La longueur d'onde (ou la température) du rayonnement subit également un effet Doppler, dû au mouvement particulier de la région d'émission dans la direction radiale à l'observation.
- Le découplage ne se produisant pas instantanément, les fluctuations sont moyennées le long de l'épaisseur de la surface de dernière diffusion, ce qui a pour effet d'atténuer les plus petites échelles de fluctuations. Le même effet d'annulation des petites échelles résulte de la dissipation thermique des oscillations, appelée amortissement de Silk.

Enfin, rappelons que si ces phénomènes décrivent des fluctuations dans  $\mathbb{R}^3$  (et même  $\mathbb{R}^4$  si l'on tient compte du temps), seules leurs traces sur une sphère (la surface de dernière diffusion) est accessible à l'observation, d'où le terme physique d'*anisotropies*.

Cette description, même sommaire, montre à quel point les propriétés statistiques des fluctuations (dont on donnera plus loin un sens plus précis) sont reliées aux propriétés de l'Univers à cette époque, c'est-à-dire aux paramètres cosmologiques. Même si ces effets sont imbriqués les uns dans les autres, les physiciens savent les quantifier. À partir de statistiques bien choisies des observations, on peut donc estimer les paramètres cosmologiques. La figure 1.3 illustre comment l'analyse statistique du Fond diffus cosmologique permet une telle estimation et comment cette estimation peut être confrontée à celles réalisées à partir d'autres types d'observations (indépendantes du *CMB*). La faible taille des régions de confiance justifie l'expression de "cosmologie de précision".

Comme expliqué dans les Chapitres 2 et 3, le spectre de puissance angulaire résume la structure de corrélation du champ. C'est généralement à partir du spectre de puissance estimé qu'on remonte aux paramètres cosmologiques. Par exemple, les deux premiers effets mentionnés ci-dessus sont responsables des pics, dit acoustiques, dans le spectre de puissance. L'existence de pics dans le spectre correspond au fait que les fluctuations sont plus marquées à certaines échelles spatiales, ce qui est un effet attendu du temps limité que les oscillations ont duré. Les échelles les plus marquées sont celles dont la durée des oscillations jusqu'au découplage est un multiple de la demi-période d'oscillation, car les fluctuations de cette taille sont, au moment du découplage, dans un état de compression, ou de décompression, maximal.

Tout cela montre l'intérêt et l'importance de l'étude statistique du Fond diffus cosmologique pour la connaissance des paramètres décrivant l'évolution de l'Univers. Même si les théories physiques prévoient des observables encore bien plus anciens (comme un Fond diffus de neutrinos ou encore des ondes gravitationnelles), c'est aujourd'hui, après les éléments légers synthétisés encore plus tôt, la plus vieille trace observée de l'histoire cosmique. Les fluctuations observées peuvent être interprétées comme les "embryons" des grandes structures, formées après le découplage par effondrement gravitationnel (que les photons ne compensent plus).



F . 1.3 – Figure de gauche : régions de confiance pour l’estimation jointe des paramètres  $\Omega_m$  et  $\Omega_\Lambda$ , à partir de l’observation du Fond diffus cosmologique (par l’expérience WMAP, en bleu) et, superposées, régions de confiance obtenues à partir d’autres observations astrophysiques. Les deux niveaux de couleurs dans chaque cas correspondent aux régions de confiance aux niveaux 95% et 68% respectivement [image : European Southern Observatory]. Figure de droite : Selon le même principe, régions de confiance au niveau 95% pour l’estimation des paramètres  $\Omega_M$  et  $H_0$  [image : Nature].

Un intérêt supplémentaire de cette étude, en dehors de la cosmologie, concerne la physique des particules. En effet, le Fond diffus révèle, bien qu’indirectement, les phénomènes physiques de l’Univers primordial, donc des phénomènes impliquant des niveaux d’énergie inatteignables sur Terre.

## Fluctuations secondaires

Les sources de fluctuations secondaires se situent entre le découplage et nous. Elles modifient les propriétés statistiques initiales du Fond diffus cosmologique. Parmi les sources identifiées, citons :

- L’interaction des photons avec les électrons libres intergalactiques. Ces derniers proviennent de la réionisation : au moment de la formation des premières étoiles, l’ensemble de la matière neutre intergalactique s’est trouvé à nouveau ionisée.
- L’effet Sunyaev-Zeldovich cinétique, due à l’effet Doppler induit par le mouvement propre des amas de galaxies au sein desquels les photons interagissent avec les électrons libres.
- L’effet Sachs-Wolfe intégré, dû à la traversée par les photons de structures en train d’évoluer. Si la structure est en train de s’effondrer, la force de gravitation est plus forte lorsque le rayonnement en ressort que lorsqu’il y est entré. Le rayonnement voit donc sa température diminuer. Si à l’inverse la structure est en train de se dissiper sous l’effet de l’expansion, la température du rayonnement qui la traverse augmente.
- Des effets de lentille gravitationnelle peuvent dévier localement la trajectoire des photons.
- Enfin, on peut classer dans cette catégorie l’effet dominant dans les fluctuations observées : la fluctuation dipolaire dû à l’effet Doppler induit par notre mouvement propre (c’est-à-dire principalement le mouvement de l’amas de galaxies auquel la Voie Lactée appartient) par rapport au référentiel du Fond diffus. Créant une augmentation de la température apparente d’environ  $3 \cdot 10^{-3}$  K

dans la direction de déplacement de notre amas et une diminution d'autant dans la direction opposée (à comparer à la moyenne 2.7 K et aux autres fluctuations de quelques  $10^{-5}$  K), cet effet fut mesuré dès 1971 [85]. Si l'on attribue la totalité de cette fluctuation à cet effet Doppler, on peut calculer notre vitesse de déplacement dans l'Univers : environ 600 km/s.

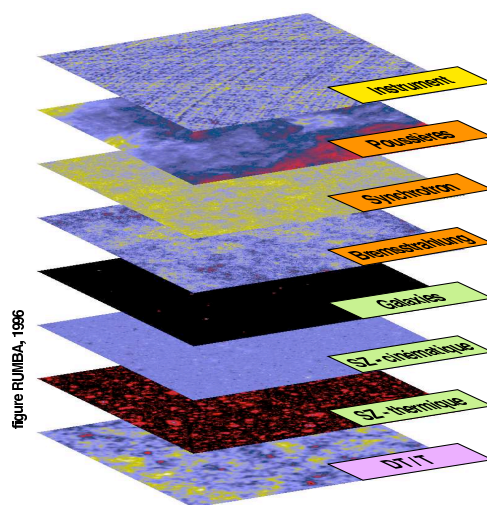
Hormis le dernier effet, qu'il faut en pratique traiter séparément (c'est-à-dire le soustraire des données, tout comme la valeur moyenne, pour pouvoir étudier les autres fluctuations) ces fluctuations secondaires présentent un grand intérêt en cosmologie. Par exemple, la mesure, via l'effet Sunyaev-Zeldovich, de la distribution spatiale et de la température des amas de galaxies renseigne sur des paramètres cosmologiques dont peu d'information est contenue dans les propriétés statistiques des fluctuations primaires.

## Avant-plans

La réalisation de cartes des fluctuations du Fond diffus cosmologique se fait par l'observation du ciel à certaines longueurs d'onde et la mesure de la puissance lumineuse reçue dans ces longueurs d'onde. Or, même dans les longueurs d'onde où le *CMB* est le plus important, d'autres objets situés plus près de nous, et qualifiés pour cela d'avant-plans, émettent de la lumière qui se superpose à celle du *CMB* et dont l'intensité est comparable à celle des fluctuations (figure 1.4). En premier lieu, l'atmosphère terrestre émet un rayonnement qui rend l'observation des fluctuations depuis le sol difficile, d'où les multiples expériences embarquées à bord de ballons stratosphériques ou de satellites. D'autres avant-plans sont, comme l'atmosphère, diffus. Ils proviennent de différents constituants de la Voie Lactée (voire la figure 1.5) :

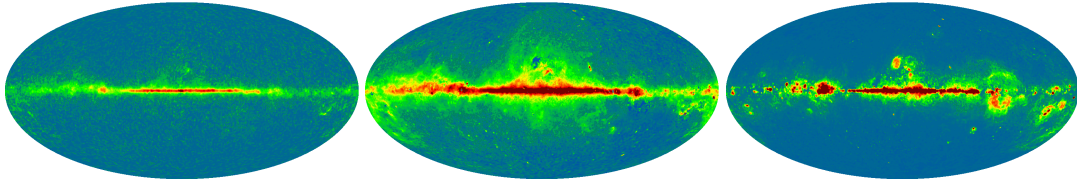
- la poussière répartie entre les étoiles qui émet thermiquement (une partie de la poussière, électrisée et en rotation, pourrait aussi émettre du fait de cette rotation),
- les électrons issus d'explosions de supernovae, accélérés dans le champ magnétique de la galaxie, émettent un rayonnement, dit synchrotron, en perdant de l'énergie,
- l'échauffement du gaz interstellaire par le rayonnement ultra-violet des étoiles produit des électrons libres qui émettent un rayonnement de freinage.

Certains avant-plans ne sont pas diffus mais sont, au contraire, des sources ponctuelles de rayonnement. Ce sont les galaxies proches, les quasars. Les amas de galaxies peuvent enfin constituer des avant-plans au travers de l'effet Sunyaev-Zeldovich thermique, dû à l'échauffement des photons par des électrons libres lorsqu'ils traversent ces amas.



F . 1.4 – Fluctuations du *CMB* (en dessous) et avant-plans [image : Bouchet et Gispert 1996].

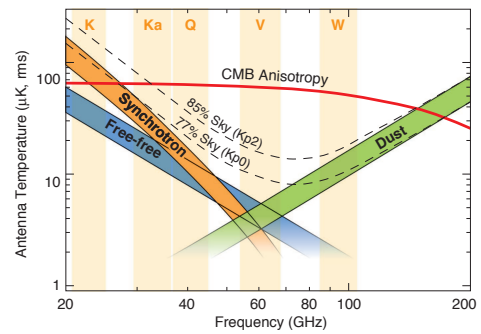




F . 1.5 – Modèles de cartes d’émission de poussière (à gauche), synchrotron (au centre) et de freinage (à droite). Ces cartes sphériques, comme toutes celles qui suivront, utilisent la projection de Mollweide [images : WMAP].

Les avant-plans n’apportent pas directement d’information cosmologique (hormis l’effet Sunyaev-Zeldovich thermique qui informe sur les amas). C’est pourquoi ils sont considérés comme une nuisance pour l’observation et l’analyse des fluctuations du *CMB* (leur étude est par ailleurs intéressante pour la connaissance, par exemple, de notre galaxie). Cependant, les rayonnements émis par ces avant-plans n’ont pas un spectre lumineux dérivé de corps noir. En dressant des cartes d’observations à plusieurs longueurs d’onde, on peut donc espérer séparer ces différentes composantes du Fond diffus (voir la section 3.3 et le chapitre 8).

Pour conclure sur les fluctuations du Fond diffus, notons qu’il existe également, en plus des fluctuations de sa température apparente, des fluctuations de sa polarisation. Ces dernières ont une amplitude encore plus faible que les fluctuations de température et leur détection reste aujourd’hui difficile. Mais leurs propriétés statistiques dépendent des paramètres cosmologiques d’une façon différente de la température et apportent donc une information complémentaire pour l’estimation de ces paramètres. C’est pourquoi un objectif important des prochaines expériences est l’observation de la polarisation.



F . 1.6 – Température équivalente (non constante car ce ne sont pas des corps noirs) en fonction de la fréquence lumineuse, pour trois les trois émissions galactiques .

### 1.3 Données disponibles et conséquences pour l'analyse statistique

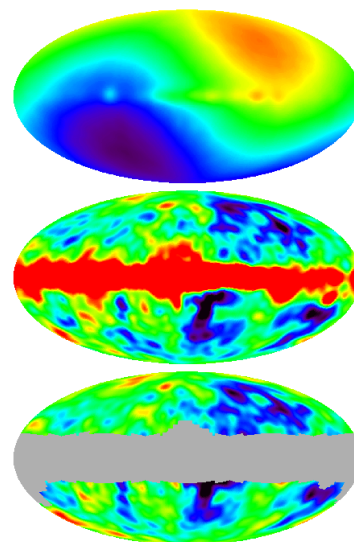
Une rapide comparaison des figures 1.7, 1.8 et 1.9 donne une idée des progrès réalisés dans l'observation des fluctuations du Fond diffus cosmologique ces deux dernières décennies. Et encore, la dernière illustration ne rend pas entièrement compte des plus de 3 millions de pixels de la carte réellement mesurée, ni de la précision accrue de la mesure sur chaque point, ni des dix cartes mesurées (dont l'illustration est une synthèse). La masse des données disponibles est rapidement devenue considérable, d'autant que le nombre d'expériences dédiées à l'observation du *CMB* a été multiplié lui aussi. Ces progrès dans l'observation vont se poursuivre dans les années à venir, augmentant d'autant la quantité et la précision des données.

La liste qui suit de quelques expériences passées et à venir, bien loin d'être exhaustive, vise à illustrer les principales caractéristiques des données disponibles et en partie leur variété et leur complexité (ainsi que les efforts d'observation réalisés). Le site dédié<sup>8</sup> de la NASA recense 36 missions ayant cartographié les fluctuations du *CMB* et 17 autres programmées dans les années à venir.

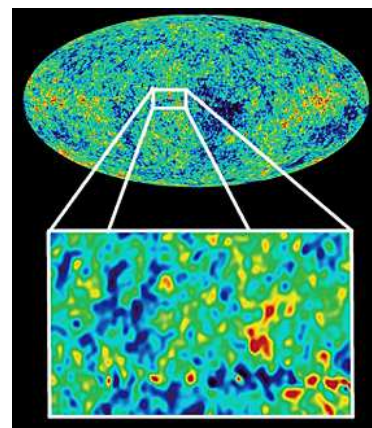
On notera en particulier que si quelques expériences (satellitaires) dressent des cartes de l'ensemble du ciel, la plupart n'en observent que certaines régions, parfois très petites. De plus, même quand la totalité du ciel a été observée, l'émission des avant-plans dans la direction de la Voie Lactée ne peut parfois pas être totalement ou suffisamment soustraite, de sorte que les données dans certaines zones sont considérées perdues (on voit clairement le bandeau galactique sur la carte de la figure 1.8).

Une autre caractéristique des données expérimentales est qu'elles sont obtenues par des balayages du ciel, où toutes les directions du ciel observées ne le sont pas le même nombre de fois. Il en résulte que l'erreur de mesure, autrement dit la variance du bruit expérimental, dépend de la direction. Elle est décrite par des *cartes de niveau de bruit*.

Troisièmement, les observations se font toujours en installant des appareils de mesure au foyer de télescopes. Chaque expérience est donc aussi caractérisée par la résolution angulaire de l'instrument. Le champ des fluctuations n'est donc pas échantillonné en un certain nombre de points mais convolué par la réponse impulsionnelle instrumentale (le *beam*, voir la figure 1.10). Les deux qualités de la réponse impulsionnelle sont sa taille (mesurée par exemple par la largeur à mi-hauteur) et sa forme. La plupart des analyses suppose une forme gaussienne bi-dimensionnelle. Cette approximation est



F . 1.8 – COBE (1992) .



F . 1.9 – WMAP (2003) .

<sup>8</sup><http://lambda.gsfc.nasa.gov>

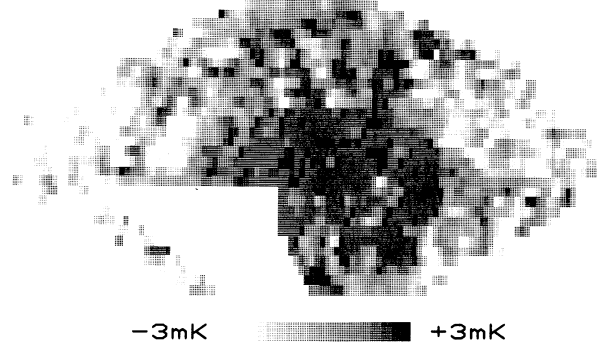


FIG. 3.—Map at 3 mm in celestial coordinates ( $\alpha, \delta$ ).  $\alpha = +24^\circ$  at left and  $0^\circ$  at right;  $\delta = +77^\circ$  at the top and  $-68^\circ$  at the bottom. Projection is  $\sin \delta$ . Pixels are nominally  $3^\circ \times 5^\circ$  ( $\alpha \times \delta$ ).  
LUBIN, VILLELA, EPSTEIN, AND SMOOT (see page L4)

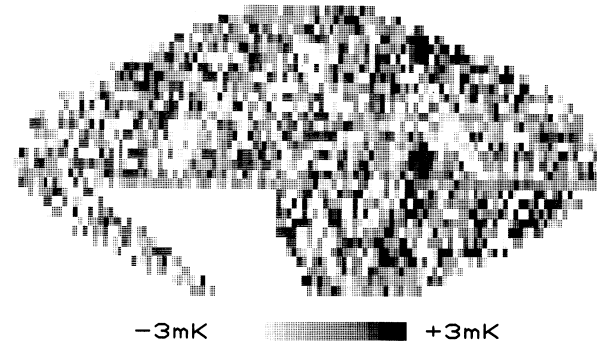


FIG. 4.—Residual map after subtraction of the dipole. Other parameters are same as Fig. 3.  
LUBIN, VILLELA, EPSTEIN, AND SMOOT (see page L4)

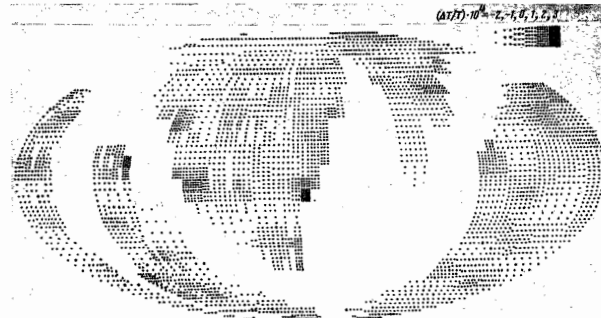


FIG. 1. A map of the temperature fluctuations in the 8-mm radiation over the celestial sphere, with the dipole component eliminated. Blank regions represent zones of interfering lunar and terrestrial radiation (right center) and a strip along the galactic equator. The measurements have been smoothed by a  $7^\circ$  Gaussian filter. The density of dots in each elementary area indicates the  $\Delta T/T$  value.

F . 1.7 – Les deux cartes du haut ont été présentées en 1985 par Lubin, Villela, Epstein et Smoot [102] suite à une série d'observations à bord de vols stratosphériques. La première inclut la composante dipolaire, qui a été soustraite de la seconde. La carte du bas est la première carte dressée à partir d'une observation spatiale en 1984 (RELIKT). Elle est tirée de l'article de Klypin, Sazhin, Strukov et Skulachev [95]. Cette dernière carte résume 15 millions de mesures moyennées (la composante dipolaire ainsi que les régions contaminées par la Voie Lactée, la Lune et la Terre ont été retirées). Cependant, même si cette carte est visiblement moins bruitée que la précédente, les inhomogénéités ne sont pas significativement supérieures à la précision des mesures.

meilleure que ne pourrait le laisser croire la figure 1.10, compte tenu du moyennage qui est opéré, pour chaque expérience et pour chaque direction, entre des observations avec différentes orientations de l'instrument.

## Exemples d'expériences passées et en cours

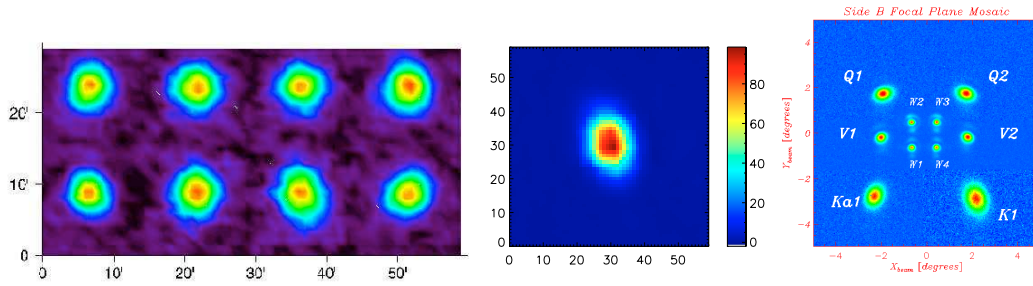
La découverte du *CMB* en 1965 a été immédiatement suivie de nombreuses tentatives d'en observer les fluctuations, à partir du sol ou de ballons (voir la figure 1.7, cartes du haut), mais seule la composante dipolaire a pu être mesurée dans les premières années.

La première expérience à annoncer la détection (à un niveau de confiance de 90%) d'autres fluctuations (la composante quadripolaire, c'est-à-dire la plus grande échelle angulaire après le dipôle) fut l'expérience soviétique RELIKT-1 en 1984, décrite par exemple dans l'article [135]. Il s'agissait de la première observation à bord d'un satellite. Un récepteur radio sensible à une bande de fréquence lumineuse autour de 37 GHz balaya le ciel pendant 6 mois. Les données manquantes et retirées représentent 30% du ciel, la résolution angulaire était d'environ  $5^\circ$  et la précision d'environ  $5 \cdot 10^{-4}$  K. Une seconde mission de bien meilleure précision et prévue pour réaliser des mesures dans 5 bandes de fréquence, a été annulée après la disparition de l'URSS.

La seconde mission satellitaire, COBE [132], lancée par la NASA, a utilisé des radiomètres différentiels (mesurant non pas des intensités – ou températures – absolues mais des différences d'intensité entre deux directions). Sa résolution angulaire (environ  $7^\circ$ ) n'était pas meilleure que celle de RELIKT-1, mais sa précision était bien supérieure (environ  $10^{-5}$  K) et 3 bandes de fréquences lumineuses étaient observées (autour de 30, 53 et 90 GHz). Outre les fluctuations du Fond diffus, les instruments de COBE ont mesuré également, avec une extrême précision, son spectre lumineux (confirmant son caractère de corps noir) ainsi que le rayonnement infrarouge de la Voie Lactée pour aider à la soustraction de cette émission aux longueurs d'onde intéressantes pour le *CMB*. En 1992, cette expérience a permis la première détection fiable de fluctuations pour toute une gamme de grandes échelles angulaires.

Au tout début des années 2000, des expériences menées à partir du sol ont permis des observations sur de petites portions du ciel, mais avec de meilleures résolutions angulaires. Ainsi ACBAR [126, 124] utilise, depuis 2001, 16 bolomètres installés derrière un télescope de 2 m au Pôle Sud (pour limiter les perturbations atmosphériques et thermiques). Les bolomètres sont répartis sur 3 bandes de fréquence lumineuse (150, 220 et 280 GHz). Ils mesurent, par leur élévation de température, la puissance lumineuse reçue à ces fréquences (donc la température du rayonnement). Pour que leur propre chaleur ne perturbe pas les mesures, ils doivent être refroidis à moins d'un tiers de degré K. Plusieurs systèmes cryogéniques imbriqués permettent de maintenir une telle température pendant 4 jours. Le niveau de bruit des mesures est estimé à environ  $2 \cdot 10^{-4}$  K /  $\sqrt{s}$ , ce qui aboutit à une précision de  $10^{-5}$  K après une saison d'observations. La résolution angulaire de l'instrument est de moins de  $0.08^\circ$ , ce qui permet l'observation de fluctuations aux petites échelles angulaires. Par contre, seules des zones du ciel de quelques degrés sont cartographiées.

ARCHEOPS et BOOMERANG ont réalisé entre 1998 et 2003 des mesures à partir de ballons volant à 40 km d'altitude. ARCHEOPS [55] cartographia une zone d'environ 25% du ciel, avec une résolution angulaire inférieure à  $0.2^\circ$  et des instruments (bolomètres) semblables à ceux qui seront utilisés pour la future mission PLANCK. Les résultats de BOOMERANG, décrit par exemple dans [66] pour le premier vol et [93] pour le plus récent, permettent une analyse à des résolutions angulaires plus fines



F . 1.10 – Exemples de réponses impulsionnelles (beams). De gauche à droite : ACBAR (diamètres d'environ  $0.1^\circ$ ), ARCHEOPS (diamètre d'environ  $0.2^\circ$ ) et WMAP (diamètres entre  $0.5^\circ$  et  $1.5^\circ$ ).

sur un carré d'environ  $10^\circ$  de côté.

WMAP est une expérience semblable à COBE (utilisation de radiomètres différentiels à bord d'un satellite) mais bien mieux résolue et plus précise. Placé en orbite solaire autour du point de Lagrange L2 (aligné avec le Soleil et la Terre) depuis 2001, WMAP continue aujourd'hui à prendre des mesures. Parmi les multiples articles décrivant ses résultats, retenons [53, 87] qui donnèrent les premières cartes et estimations du spectre après un an de prise de données, et [88, 114] qui fournissent les résultats les plus récents. Dix couples d'instruments, répartis dans 5 bandes de fréquences lumineuses entre 23 et 94 GHz, réalisent des cartes du ciel. La précision des mesures  $\sigma_0$  est estimée entre  $1.4 \cdot 10^{-3}$  et  $7 \cdot 10^{-3}$  K, selon les bandes de fréquences, pour chaque observation dans une direction. Le nombre d'observations  $N_{obs}$  n'est pas le même pour toutes les directions : il dépend de la stratégie de balayage du ciel adoptée, dont la figure 1.11 donne une idée. Ce nombre, au bout de 5 ans, varie selon les directions entre 1000 et 8000, ce qui donne un niveau de bruit  $\sigma_0 N_{obs}^{-1/2}$  pour les mesures entre  $1.5 \cdot 10^{-5}$  et  $2.3 \cdot 10^{-4}$  K. L'expérience WMAP est à ce jour celle qui fournit, après une analyse appropriée pour notamment soustraire les avant-plans (voir le chapitre 8), la meilleure carte des fluctuations du *CMB* sur l'ensemble du ciel, des plus grandes échelles jusqu'à celles déterminées par la résolution des instruments, soit moins de  $0.3^\circ$ .

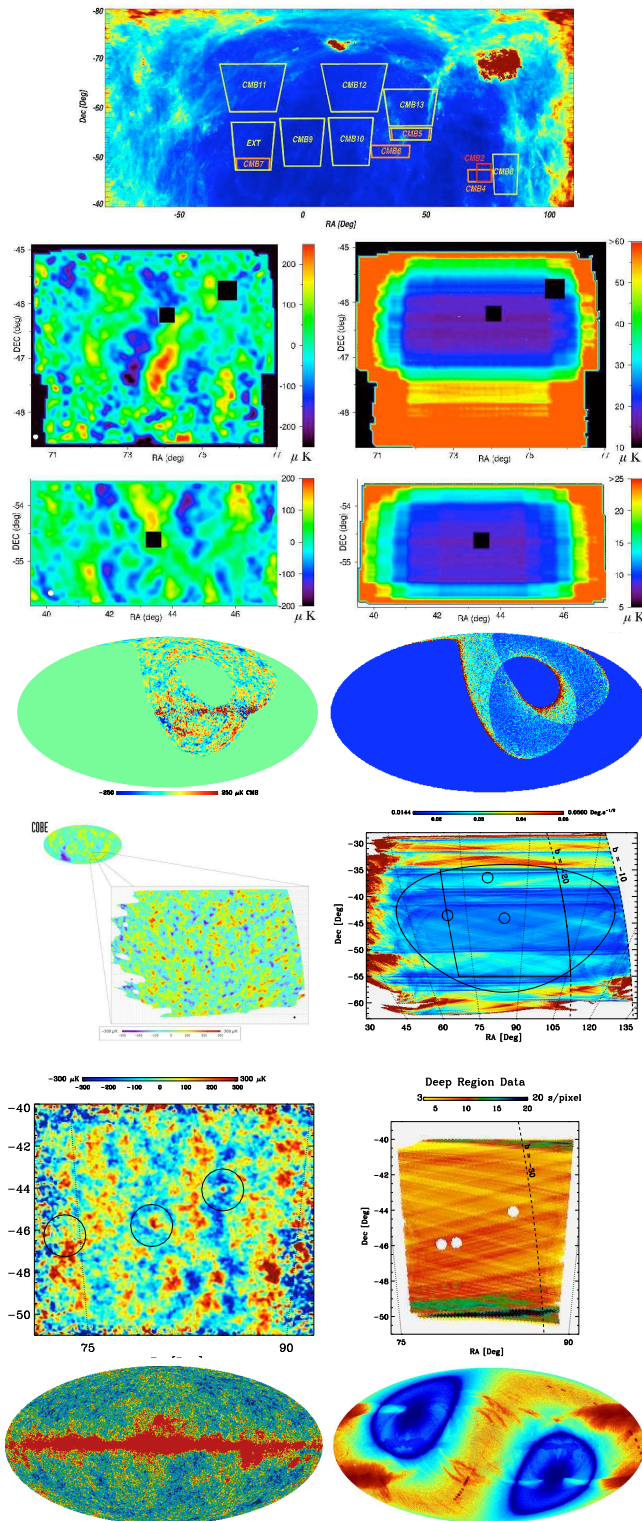
ARCHEOPS, BOOMERANG et WMAP, outre la température du rayonnement, ont également mesuré sa polarisation, mais pour cette dernière, la précision actuelle des expériences reste insuffisante. Les prochaines missions d'observation auront pour principal objectif d'améliorer cette mesure.

La figure 1.11 illustre le type de cartes dressées par ces expériences ainsi que les cartes de niveau de bruit correspondantes. La figure 1.10 montre des exemples de réponses impulsionnelles instrumentales.

## Exemples d'expériences programmées pour un avenir proche

PLANCK [99, 141] est une mission de l'Agence Spatiale Européenne, à bord d'un satellite qui devrait être lancé début 2009. La température du rayonnement sera mesurée dans dix bandes de fréquences lumineuses, entre 30 et 900 GHz, par 104 détecteurs au dos d'un télescope de 1,5 m. La résolution angulaire sera comprise entre  $0.08^\circ$  et  $0.5^\circ$  et la précision de mesure entre  $7 \cdot 10^{-3}$  et  $2 \cdot 10^{-6}$  K, selon les bandes. PLANCK a pour objectif de fournir une carte des fluctuations du *CMB* à une précision telle que l'estimation du spectre de puissance atteigne la variance du cas non bruité (dite variance cosmique, voire le chapitre 3). Les physiciens espèrent également obtenir de cette expérience une





Vue d'ensemble des zones observées par ACBAR (en fond : une carte de poussière)

Détail d'une zone observée par ACBAR (notée CMB2 ci-dessus)

Une autre zone observée par ACBAR (CMB5)

ARCHEOPS

BOOMERANG, vol de 1998

BOOMERANG, vol de 2003, zone "Deep"

WMAP (bande de fréquence dite "Q")

F . 1.11 – Exemples de zones observées et de niveaux de bruit pour des expériences passées. A gauche : observations. A droite : cartes de niveau de bruit (ou à l'inverse, selon les cas, cartes de nombre d'observations). [images : collaborations des expériences].

carte relativement fiable des fluctuations de polarisation du Fond diffus.

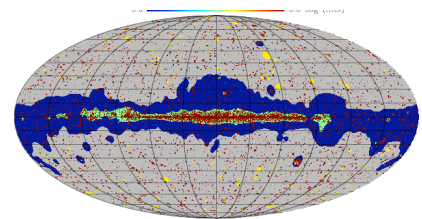
D'autres expériences fourniront des cartes de petites zones de meilleure résolution. ACT [96] par exemple, qui utilisera plus de 3000 bolomètres derrière un télescope de 6 m dans le désert d'Atacama, devrait fournir des cartes d'une bande circulaire de  $2^\circ$  de largeur à une résolution de  $0.02^\circ$ .

D'autres expériences encore, comme EBEX [116] ou QUIET [129] seront spécialement dédiées à la mesure des fluctuations de polarisation. QUIET prendra des mesures depuis un télescope de 7 m au sol, et EBEX grâce à un télescope d'1,5 m emporté par un ballon, les deux expériences utilisant plusieurs centaines de bolomètres. La qualité (résolution et précision) de ces mesures de polarisation sur des petites zones (4 carrés de  $6^\circ$  de côté pour QUIET) sera bien meilleure que celle de PLANCK.

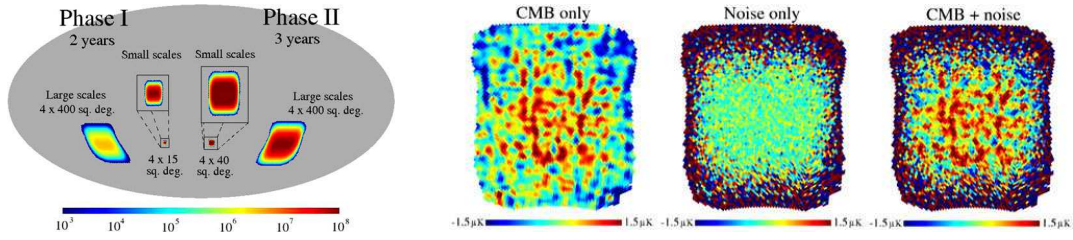
## Conséquences pour l'analyse statistique

Au vu de ces quelques exemples, revenons sur certaines particularités des cartes de fluctuations du *CMB*. Tout d'abord, la *masse de données* collectée croît à un rythme élevé. Cette constatation est commune à de nombreux domaines scientifiques, mais elle est particulièrement vraie ici. Chacune des expériences à venir (ACT, QUIET, PLANCK, etc.) annonce plusieurs dizaines de terabits de données collectées par année d'observation, ce qui représente un défi pour leur traitement, d'un point de vue technologique mais aussi statistique, pour être à la hauteur de l'amélioration de la qualité des observations. Cet accroissement de la quantité de données traduit en effet une amélioration dans deux directions : la précision des mesures et la résolution angulaire. L'incidence de la première sur l'analyse est que des phénomènes (par exemple des avant-plans) qui pouvaient être tenus pour négligeables au regard des anciennes précisions ne le peuvent plus. Autrement dit, avec la précision augmentée aussi la complexité des données. Quant à la meilleure résolution angulaire, elle permet la production de cartes de plus en plus volumineuses (en nombre de pixels  $N_{pix}$ ) et une analyse des fréquences spatiales plus élevées. Ainsi, COBE a fourni 3 cartes de quelques milliers de pixels, WMAP 10 cartes de 3 millions de pixels et PLANCK devrait fournir une centaine de cartes de plus de 50 millions de pixels. La récente mise à jour du logiciel d'analyse sur la sphère HEALPIX a augmenté le nombre limite de pixels analysables à  $3 \cdot 10^{18}$ .

Par ailleurs, la zone du ciel couverte par chaque expérience est variable en taille et en localisation. Peu d'expériences couvrent la totalité de la sphère céleste, et même lorsque c'est le cas, des *masques* sont souvent appliqués sur les données, de sorte que le problème de la gestion de *données manquantes* se pose presque systématiquement. Ces données manquantes empêchent toute utilisation directe des techniques de Fourier. Or ces dernières sont d'autant plus indispensables que la résolution angulaire est fine : sous l'hypothèse de stationnarité, la compression d'information apportée par le passage dans l'espace de Fourier grandit quand l'échelle angulaire rétrécit. Les techniques requérant l'utilisation de matrices de covariances dans le domaine direct ( $N_{pix} \times N_{pix}$ ) sont en pratique hors de portée pour des cartes à haute résolution. Deux types de données manquantes sont à envisager : de larges zones connexes de la sphère (zones non observées et/ou corrompues par des avant-plans diffus) et de multiples petites zones réparties sur toute la sphère (avant-plans localisés). Ces deux situations sont illustrées par la figure 1.12.



F . 1.12 – Exemples de masques (re-pérés par leur couleur) utilisés dans l'analyse de données simulées de PLANCK [image : [100]].

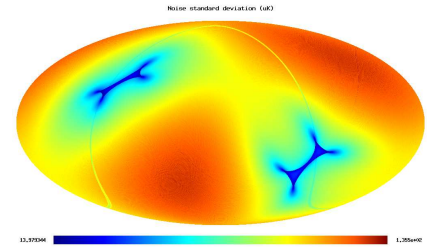


F . 1.14 – A gauche : zones d’observation (cartes de nombre d’observations) prévues pour QUIET. A droite : simulation d’observations dans une petite zone. [images : QUIET]

Une autre propriété inhérente à la façon dont les données sont obtenues est le caractère *hétéroscédastique* du bruit de mesure, clairement illustré par les cartes de niveau de bruit. Toute analyse fine des données doit donc tenir compte de cette inhomogénéité de la précision.

Comme on l’a vu, l’observation par des instruments ayant une réponse impulsionnelle finie génère un problème inverse, plus précisément de *déconvolution* sphérique, avec un “beam” supposé à symétrie circulaire ou, si possible, plus réaliste.

Il ressort enfin clairement qu’il y a potentiellement beaucoup à gagner de l’utilisation *conjointe* des données fournies par les diverses expériences, puisque celles-ci sont complémentaires. Or ceci ne va pas sans poser des difficultés. Une première, d’ordre technique, est que le schéma de pixelisation adopté pour représenter les observations n’est pas nécessairement le même pour toutes les expériences. Se pose alors le problème de la “repixelisation”, qui doit être réalisée de façon à engendrer le moins de biais possible et perdre le moins d’information (quant aux données “brutes”, non-pixelisées, fournies par les instruments, elle sont généralement assez complexes pour n’être utilisables de façon fiable que par ceux qui les ont produites). Un autre enjeu, plus fondamental, est la possibilité d’agrégier des expériences en tenant compte de leur complexité, c’est-à-dire du fait que les zones couvertes sont différentes et se recouvrent partiellement, que les niveaux de bruit et surtout les beams sont différents d’une expérience à l’autre, etc.



F . 1.13 – Carte de niveau de bruit de PLANCK (pour une hypothèse plausible de balayage du ciel).

## Annexe

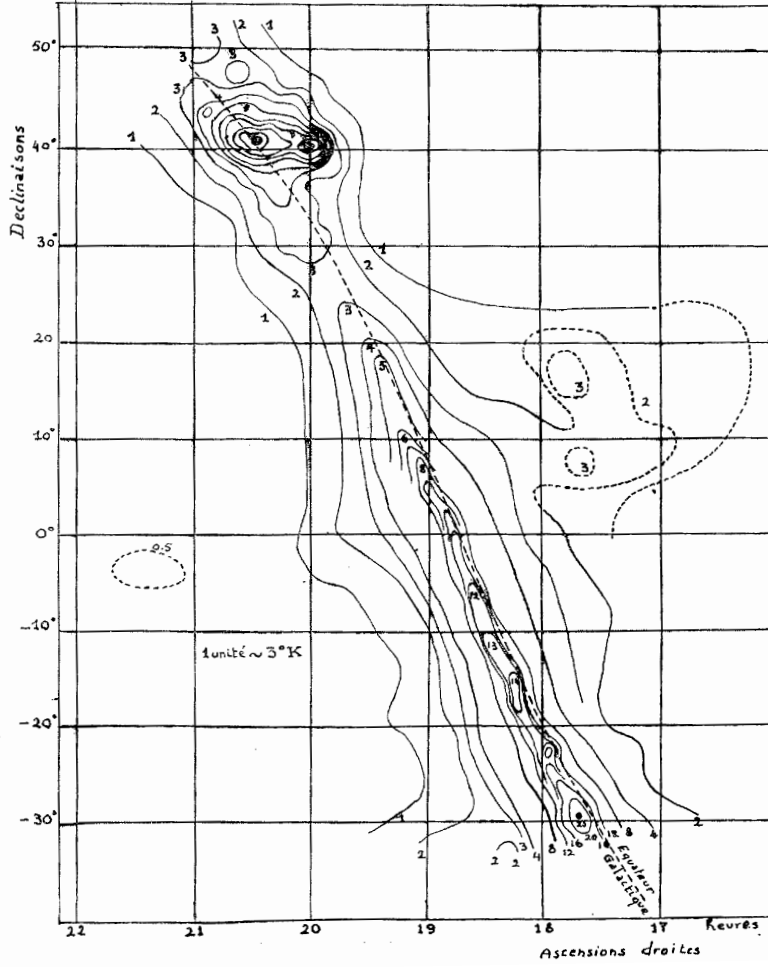
Esquissons ci-dessous le raisonnement concernant l’évolution d’une distribution de corps noir dans un univers en expansion (page 13). Soit  $\lambda_0 \mapsto u_{T_0}(\lambda_0)$  la distribution, selon les longueurs d’onde, de l’énergie lumineuse par unité de volume d’un corps noir de température  $T_0$  à un instant  $t_0$ . Par définition,

$$B(\lambda_0, T_0) = \frac{du_{T_0}}{d\lambda_0}(\lambda_0) = \frac{2hc^2}{\lambda_0^5} \left( \exp\left(\frac{hc}{kT_0\lambda_0}\right) - 1 \right)^{-1}.$$

Par ailleurs, l’énergie d’un photon de longueur d’onde  $\lambda_0$  est  $E = \frac{hc}{\lambda_0}$ .

Dans un univers en expansion d’un facteur  $a := \frac{a(t)}{a(t_0)}$  entre les instants  $t_0$  et  $t$ , les distances sont multipliées par  $a$  et les volumes par  $a^3$ . Donc





F . 1.15 – Carte de la température du rayonnement céleste au voisinage de la Voie Lactée présentée par Denisse, Le Roux et Steinberg [70] en 1955. En dehors du plan galactique, cette carte reporte un rayonnement à une température d'environ 3 K.

- la densité de photon est multipliée par  $a^{-3}$ ,
  - la longueur d'onde de chaque photon est multipliée par  $a$ ,
- donc la densité volumique d'énergie est multipliée par  $a^{-4}$ .

Ainsi, en notant  $\tilde{u}$  la nouvelle distribution de densité volumique d'énergie par longueur d'onde à l'instant  $t$ ,

$$\tilde{u}(\lambda) = a^{-4} u_{T_0}(a^{-1} \lambda)$$

et

$$\frac{d\tilde{u}}{d\lambda}(\lambda) = a^{-5} B(a^{-1} \lambda, T_0) = B(\lambda, \frac{T_0}{a}).$$

Autrement dit, la distribution  $\tilde{u}$  est encore celle d'un corps noir, de température  $\frac{T_0}{a}$ .

## Chapitre 2

# Analyse sur la sphère

### Sommaire

2.1	Analyse harmonique dans $\mathbb{L}^2(\mathbb{S}^2, d\xi)$ . . . . .	<b>30</b>
	Mesure de Lebesgue et Laplacien sphériques . . . . .	30
	Harmoniques sphériques . . . . .	32
	Base usuelle des harmoniques sphériques . . . . .	34
	Construction à partir de la théorie des groupes . . . . .	34
2.2	Champs aléatoires sur la sphère . . . . .	<b>35</b>
	Champs aléatoires et distributions directionnelles . . . . .	35
	Champs aléatoires stationnaires sur la sphère . . . . .	36
	Exemples d'applications . . . . .	38
2.3	Ondelettes sphériques . . . . .	<b>39</b>
	Motivations . . . . .	39
	Premiers travaux . . . . .	41
	Les needlets : propriétés de localisation et propriétés statistiques . . . . .	44

L'étude du Fond diffus cosmologique introduite au chapitre précédent conduit à l'analyse de fonctions, déterministes ou aléatoires, définies sur l'ensemble des directions de l'espace, cet ensemble pouvant être identifié à la sphère. L'étude de fonctions définies sur la sphère ou à valeurs sphériques n'est cependant pas limitée à cette application : on rencontre de telles fonctions dans de nombreux domaines scientifiques.

Historiquement, les principaux domaines concernés sont l'astronomie et la géophysique (en particulier, pour cette dernière, l'étude du champ de gravitation et du champ magnétique terrestres, et celle des ondes sismiques) et ces deux disciplines ont encore joué ces dernières années un rôle moteur dans l'accumulation de données sphériques et le développement d'outils théoriques pour leur analyse. Les champs de température et de pression en météorologie, les courants océaniques, les déformations organiques en imagerie médicale, les données directionnelles (par exemple de déplacement ou d'audition) issues d'études comportementales (humaines et animales), le mouvement des pôles en paléomagnétisme, les objets tridimensionnels en vision artificielle, les positionnements par satellites (GPS) – fournissent autant d'exemples supplémentaires de données sphériques.

Les multiples problèmes mathématiques et statistiques posés par l'analyse de données sphériques (ceux présentés en détails dans le Chapitre 3 à propos du *CMB*, mais aussi la restauration d'images,

l'estimation de densité, etc.) sont actuellement l'objet d'une activité intense. Au moins quatre numéros de revues ont été entièrement consacrés à ces questions ces cinq dernières années [1, 2, 35, 36]. Les problèmes équivalents dans les espaces euclidiens usuels  $(\mathbb{R}, \mathbb{R}^n)$  sont connus de longue date et des techniques efficaces ont été développées et continuent de l'être, telles les ondelettes. La généralisation de ces techniques au cas de la sphère ou à celui d'autres variétés est en revanche beaucoup plus récente.

De nombreux résultats exposés dans ce chapitre sont généralisables à des sphères de dimension supérieure et à des variétés riemanniennes plus générales. Cependant, seul le cas de la sphère de dimension 2 sera présenté, afin de ne pas introduire de notations et de difficultés techniques qui ne seraient plus utiles par la suite.

## 2.1 Analyse harmonique dans $\mathbb{L}^2(\mathbb{S}^2, d\xi)$

L'analyse harmonique de fonctions à valeurs sphériques étant moins commune que celle des fonctions sur les espaces euclidiens usuels, il est utile d'en rappeler les résultats de base, dont on peut trouver les démonstrations dans la littérature classique, par exemple [31, 40], [27, chap. 4 à 7], [43, chap. 4], [14, chap. 2 et 3]. Munissons l'espace  $\mathbb{R}^3$  des coordonnées sphériques  $(r, \theta, \varphi)$  où  $r > 0$  est la distance à l'origine,  $\theta \in [0, \pi]$  est la colatitude et  $\varphi \in [0, 2\pi[$  est la longitude. La sphère unité  $\mathbb{S}^2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}$ , où  $\|\cdot\|$  est la norme euclidienne usuelle, peut être identifiée à l'ensemble  $\{(\theta, \varphi) : \theta \in [0, \pi], \varphi \in [0, 2\pi[ \}$ . L'élément générique de  $\mathbb{S}^2$ , considéré comme une direction  $(\theta, \varphi)$  ou comme un vecteur de norme unité, sera noté  $\xi$ . La distance géodésique entre deux points de  $\mathbb{S}^2$ , qui est égale à l'angle entre les deux directions, est notée  $d$  :

$$d(\xi, \xi') := \arccos(\xi \cdot \xi')$$

où  $\cdot$  est le produit scalaire usuel de  $\mathbb{R}^3$ .

### Mesure de Lebesgue et Laplacien sphériques

Dans l'espace  $\mathbb{R}^n$  muni de la tribu borélienne, la mesure de Lebesgue  $dx$  est l'unique mesure (à une constante multiplicative près) invariante par translation : pour tout  $x_0 \in \mathbb{R}^n$  et toute fonction  $f$  mesurable sur  $\mathbb{R}^n$ ,  $\int_{\mathbb{R}^n} f(x + x_0) dx = \int_{\mathbb{R}^n} f(x) dx$ .

De même [11], la *mesure de Lebesgue* sur  $\mathbb{S}^2$ , notée  $d\xi$ , est l'unique mesure sur les boréliens de  $\mathbb{S}^2$  de masse totale  $4\pi$  et invariante par rotation : pour toute rotation  $\rho \in SO(3)$  et toute fonction  $f$  mesurable sur  $\mathbb{S}^2$ ,

$$\int_{\mathbb{S}^2} f \circ \rho(\xi) d\xi = \int_{\mathbb{S}^2} f(\xi) d\xi \quad .$$

Cette mesure s'exprime en coordonnées sphériques par  $d\xi = \sin\theta d\theta d\varphi$ .

Muni du produit scalaire  $\langle f, g \rangle := \int_{\mathbb{S}^2} f(\xi) g^*(\xi) d\xi$ , où  $z^*$  désigne le complexe conjugué de  $z \in \mathbb{C}$ , l'espace  $\mathbb{L}^2(\mathbb{S}^2) := \mathbb{L}^2(\mathbb{S}^2, d\xi)$  est un espace de Hilbert, dont la norme est notée  $\|\cdot\|_2$

Alternativement,  $d\xi$  peut être définie par la relation  $dx = d\xi r^2 dr$ . Plus précisément,  $d\xi$  est l'unique

mesure sur  $\mathbb{S}^2$  telle que pour toute fonction  $f$  mesurable sur  $\mathbb{R}^3$ ,

$$\int_{\mathbb{R}^3} f(x) dx = \int_0^\infty \int_{\mathbb{S}^2} f(r\xi) d\xi r^2 dr \quad .$$

L'analyse harmonique est née du projet de Fourier de comprendre et de décrire les fonctions intervenant dans la diffusion et l'équilibre de la chaleur. La séparation des variables de temps et d'espace dans l'équation de la chaleur le conduisit à considérer les fonctions propres du Laplacien, qui sont les fonctions trigonométriques. Pour développer les idées de Fourier sur des variétés riemanniennes compactes, l'opérateur de Laplace-Beltrami, qui généralise le Laplacien, est nécessaire.

Sur la sphère, on peut le construire de la façon suivante : toute fonction  $f$  définie sur  $\mathbb{S}^2$  peut être étendue en une fonction  $\tilde{f} : x = (r, \theta, \varphi) \mapsto f(\|x\|^{-1}x) = f(\theta, \varphi)$  définie sur  $\mathbb{R}^3 \setminus \{0\}$ . Lorsque cela est défini, posons

$$\Delta_{\mathbb{S}^2} f := (\Delta \tilde{f})|_{\mathbb{S}^2} \quad ,$$

où  $\Delta$  est le Laplacien dans  $\mathbb{R}^3$ .

A nouveau,  $\Delta_{\mathbb{S}^2}$  peut être défini de façon équivalente par la relation

$$\Delta = \frac{1}{r^2} \left\{ \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \Delta_{\mathbb{S}^2} \right\}. \quad (2.1)$$

L'expression explicite de l'opérateur de Laplace-Beltrami en coordonnées sphériques est

$$\Delta_{\mathbb{S}^2} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}.$$

Une de ses premières propriétés est qu'il commute avec les rotations :

**Lemme 2.1.** *Pour toute fonction  $f$  de classe  $C^2$  sur la sphère et pour toute rotation  $\rho \in SO(3)$ ,*

$$\Delta_{\mathbb{S}^2}(f \circ \rho) = (\Delta_{\mathbb{S}^2} f) \circ \rho \quad .$$

Une application du théorème de Green-Ostrogradski permet d'établir que l'opérateur  $\Delta_{\mathbb{S}^2}$  est de plus symétrique.

**Lemme 2.2.** *Pour toutes fonctions  $f$  et  $g$  de classe  $C^2$  sur la sphère,*

$$\langle \Delta_{\mathbb{S}^2} f, g \rangle = \langle f, \Delta_{\mathbb{S}^2} g \rangle \quad .$$

**Corollaire 2.1.** *Si deux fonctions de classe  $C^2$  sur la sphère  $f$  et  $g$  sont des vecteurs propres de  $\Delta_{\mathbb{S}^2}$  associés à des valeurs propres distinctes, alors  $f$  et  $g$  sont orthogonales dans  $\mathbb{L}^2(\mathbb{S}^2)$ .*

Ces vecteurs propres peuvent être reliés aux fonctions harmoniques homogènes. Une fonction  $f$  définie sur  $\mathbb{R}^3 \setminus \{0\}$  est dite homogène de degré  $\ell$  si pour tout  $t > 0$ ,  $f(tx) = t^\ell f(x)$ , c'est-à-dire si elle est de la forme  $f(x) = r^\ell f|_{\mathbb{S}^2}(\theta, \varphi)$ . Le lemme suivant est alors une conséquence immédiate de la relation (2.1).

**Lemme 2.3.** *Si  $f$  est la restriction à  $\mathbb{S}^2$  d'une fonction définie sur  $\mathbb{R}^3 \setminus \{0\}$ , homogène de degré  $\ell$  et harmonique, alors*

$$\Delta_{\mathbb{S}^2} f = -\ell(\ell + 1)f \quad .$$

*Réciproquement, toute fonction définie sur  $\mathbb{R}^3 \setminus \{0\}$ , homogène de degré  $\ell$  et dont la restriction à  $\mathbb{S}^2$  est vecteur propre de  $\Delta_{\mathbb{S}^2}$  pour la valeur propre  $-\ell(\ell + 1)$ , est harmonique.*

## Harmoniques sphériques

L'analyse harmonique sur la sphère, interprétée comme l'analyse des fréquences spatiales sphériques des fonctions, repose sur les sous-espaces propres de  $\Delta_{\mathbb{S}^2}$ , c'est-à-dire sur les harmoniques sphériques.

**Définition 2.1.** Une harmonique sphérique de fréquence  $\ell \in \mathbb{N}$  est la restriction à  $\mathbb{S}^2$  d'un polynôme homogène de degré  $\ell$ , harmonique. L'ensemble des harmoniques sphériques de fréquence  $\ell$  est noté  $\mathbb{H}_\ell$ .

D'après le lemme 2.3, l'ensemble des harmoniques sphériques est l'intersection des restrictions à  $\mathbb{S}^2$  des polynômes homogènes à trois variables avec le sous-espace propre de  $\Delta_{\mathbb{S}^2}$  associé à la valeur propre  $-\ell(\ell + 1)$ . Par ailleurs, puisque les harmoniques sphériques sont restrictions de polynômes homogènes, elles sont de même parité que leur fréquence :  $f(-\xi) = (-1)^\ell f(\xi)$  si  $f \in \mathbb{H}_\ell$ .

Examinons l'analogie des harmoniques sphériques pour les fonctions  $2\pi$ -périodiques, autrement dit les fonctions du cercle  $\mathbb{S}^1 := \{x \in \mathbb{R}^2 : \|x\|^2 = 1\}$ . En dimension 2, le Laplacien vérifie, en coordonnées polaires,  $\Delta = \frac{1}{r^2} \left\{ r \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \right) + \frac{\partial^2}{\partial \theta^2} \right\}$ . Les restrictions de fonctions homogènes de degré  $\ell$  et harmoniques sont donc vecteurs propres de  $\frac{\partial^2}{\partial \theta^2}$  (qui est l'opérateur de Laplace-Beltrami sur le cercle) pour la valeur propre  $-\ell^2$ . Elles forment donc un sous-espace de  $\mathbb{L}^2(\mathbb{S}^1, d\theta)$  de dimension 2 (sauf pour  $\ell \neq 0$ , où la dimension vaut 1) engendré, par exemple, par la base de Fourier usuelle  $e^{i\ell\theta}$  et (pour  $\ell \neq 0$ )  $e^{-i\ell\theta}$  (qui sont des restrictions de polynômes homogènes de degré  $\ell$  en les variables  $x_1 = r \cos \theta$  et  $x_2 = r \sin \theta$ ).

Revenant au cas de la sphère, le théorème suivant résume les propriétés de la décomposition d'une fonction de  $\mathbb{L}^2(\mathbb{S}^2)$  en harmoniques sphériques.

**Théorème 2.1.** 1.  $\mathbb{H}_\ell$  est un sous-espace vectoriel de  $\mathbb{L}^2(\mathbb{S}^2)$  invariant par rotation :  $\forall \ell \in \mathbb{N}$ ,  $\forall f \in \mathbb{H}_\ell, \forall \rho \in SO(3), f \circ \rho \in \mathbb{H}_\ell$ .

2.  $\dim(\mathbb{H}_\ell) = 2\ell + 1$ .

3.  $\mathbb{L}^2(\mathbb{S}^2) = \bigoplus_{\ell \in \mathbb{N}}^\perp \mathbb{H}_\ell$ .

Les éléments d'une base orthonormée de  $\mathbb{H}_\ell$  seront indexés par les entiers  $m \in \{-\ell, \dots, \ell\}$ . Soit  $(\mathcal{Y}_{\ell,m})_{-\ell \leq m \leq \ell}$  une telle base orthonormée. La projection orthogonale  $\Pi_\ell$  de  $\mathbb{L}^2(\mathbb{S}^2)$  sur  $\mathbb{H}_\ell$ , qui ne dépend pas de la base particulière choisie, peut s'écrire à l'aide du noyau reproduisant

$$\mathcal{L}_\ell(\xi, \xi') := \sum_{m=-\ell}^{\ell} \mathcal{Y}_{\ell,m}(\xi) \mathcal{Y}_{\ell,m}^*(\xi'). \quad (2.2)$$

En effet,

$$\begin{aligned} \forall f \in \mathbb{L}^2(\mathbb{S}^2), \Pi_\ell f &= \sum_{m=-\ell}^{\ell} \langle f, \mathcal{Y}_{\ell,m} \rangle \mathcal{Y}_{\ell,m} \\ &= \int_{\mathbb{S}^2} f(\xi) \mathcal{L}_\ell(\cdot, \xi) d\xi \end{aligned}$$

où  $\mathcal{L}_\ell$  vérifie la propriété

$$\forall \xi_1, \xi_2 \in \mathbb{S}^2, \int_{\mathbb{S}^2} \mathcal{L}_\ell(\xi_1, \xi) \mathcal{L}_\ell(\xi, \xi_2) d\xi = \mathcal{L}_\ell(\xi_1, \xi_2).$$

**Théorème 2.2.**  $\mathcal{L}_\ell(\xi, \xi') = L_\ell(\xi \cdot \xi')$  où  $L_\ell$  est le polynôme de Legendre de degré  $\ell$ , c'est-à-dire le polynôme de degré  $\ell$  dans la suite des polynômes orthogonaux sur  $[-1, 1]$  pour le poids 1, normalisé par  $\int_{-1}^1 L_\ell^2 = \frac{2\ell+1}{8\pi^2}$ .

Étant donnée une base orthonormée  $\mathcal{Y}_{\ell,m}$ , on note  $a_{\ell,m} := \langle f, \mathcal{Y}_{\ell,m} \rangle$  (ou  $a_{\ell,m}^f$  pour lever d'éventuelles ambiguïtés) les coefficients d'une fonction  $f$  sur cette base, appelés coefficients de Fourier. On note encore  $\widetilde{\ell}^2$  l'ensemble des suites complexes  $(a_{\ell,m})_{\ell \in \mathbb{N}, m \in \{-\ell, \dots, \ell\}}$  vérifiant

$$\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |a_{\ell,m}|^2 < \infty .$$

On appelle enfin transformation en harmoniques sphériques, ou transformation de Fourier, l'application SHT (pour Spherical Harmonic Transform)

$$\text{SHT} : f \in \mathbb{L}^2(\mathbb{S}^2) \mapsto (a_{\ell,m})_{\ell \in \mathbb{N}, m \in \{-\ell, \dots, \ell\}} \in \widetilde{\ell}^2$$

et transformation de Fourier inverse l'application

$$\text{SHT}^{-1} : (a_{\ell,m})_{\ell \in \mathbb{N}, m \in \{-\ell, \dots, \ell\}} \in \widetilde{\ell}^2 \mapsto \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell,m} \mathcal{Y}_{\ell,m} \in \mathbb{L}^2(\mathbb{S}^2)$$

vérifiant  $\text{SHT} \circ \text{SHT}^{-1} = \text{Id}_{\widetilde{\ell}^2}$  et  $\text{SHT}^{-1} \circ \text{SHT} = \text{Id}_{\mathbb{L}^2(\mathbb{S}^2)}$ .

Généralisons maintenant le noyau reproduisant défini par l'équation (2.2). Pour cela identifions une fonction  $\Lambda \in \mathbb{L}^2([-1, 1], dx)$  avec le noyau associé  $(\xi, \xi') \in (\mathbb{S}^2)^2 \mapsto \Lambda(\xi \cdot \xi') \in \mathbb{R}$ .

**Définition 2.2.** Une fonction  $R \in \mathbb{L}^2(\mathbb{S}^2)$  est dite radiale (ou à symétrie axiale, ou encore zonale) centrée en  $\xi_0 \in \mathbb{S}^2$  s'il existe  $\Lambda \in \mathbb{L}^2([-1, 1])$  telle que  $R(\xi) = \Lambda(\xi \cdot \xi_0)$ .

**Définition 2.3.** La convolution de  $f \in \mathbb{L}^2(\mathbb{S}^2)$  par une fonction  $\Lambda \in \mathbb{L}^2([-1, 1])$  est définie par

$$(\Lambda * f)(\xi) := \int_{\mathbb{S}^2} \Lambda(\xi \cdot \xi') f(\xi') d\xi' .$$

**Définition 2.4.** Les coefficients de Legendre de  $\Lambda \in \mathbb{L}^2([-1, 1])$  sont

$$\Lambda_\ell^\wedge := \frac{8\pi^2}{2\ell+1} \int_{-1}^1 \Lambda(x) L_\ell(x) dx, \quad \ell \in \mathbb{N}.$$

Par orthogonalité des polynômes de Legendre,  $\Lambda = \sum_{\ell=0}^{\infty} \Lambda_\ell^\wedge L_\ell$ . Le théorème suivant permet de traiter la convolution dans le domaine de Fourier en exprimant l'analogie d'une propriété importante de la transformée de Fourier usuelle : la convolution dans le domaine direct correspond à une multiplication dans le domaine de Fourier

**Théorème 2.3** (Formule de Funck-Hecke). Soient  $\Lambda \in \mathbb{L}^2([-1, 1], dx)$  et  $f \in \mathbb{H}_\ell$ . Alors,

$$\Lambda * f = \Lambda_\ell^\wedge f .$$

En particulier, pour toute fonction  $f \in \mathbb{L}^2(\mathbb{S}^2)$ ,

$$a_{\ell,m}^{\Lambda * f} = \Lambda_\ell^\wedge a_{\ell,m}^f .$$

## Base usuelle des harmoniques sphériques

La base orthonormée usuelle de l'espace  $\mathbb{H}_\ell$  est obtenue en séparant les variables  $\theta$  et  $\varphi$  dans l'équation  $\Delta_{\mathbb{S}^2} f = -\ell(\ell+1)f$ . Si l'on impose à la fonction  $f(\theta, \varphi)$  d'être de la forme  $f(\theta, \varphi) = a(\theta)b(\varphi)$  où  $b$  est  $2\pi$ -périodique, alors  $b$  vérifie  $b'' = -m^2b$  pour un certain entier  $m$ . Donc  $b$  est combinaison linéaire des  $e^{\pm im\varphi}$ , et  $f$  est, comme  $b$ , vecteur propre de l'opérateur  $\frac{\partial^2}{\partial \varphi^2}$ .

Comme on le voit, autant la définition de l'espace des harmoniques sphériques est indépendante du système de coordonnées sphériques de départ, autant les fonctions découlant du choix ci-dessus dépendent de ces coordonnées. L'opérateur  $\frac{\partial^2}{\partial \varphi^2}$ , par exemple, ne commute visiblement pas avec toutes les rotations.

L'équation différentielle vérifiée par la fonction  $a$ , appelée équation de Legendre associée (à l'équation de Laplace pour des variables séparées) a pour solutions les fonctions  $L_{\ell,m}(\cos \theta)$  où  $L_{\ell,m}$  sont les *fonctions de Legendre associées*. Ces dernières peuvent être définies, pour  $\ell \in \mathbb{N}$  et  $m \in \{0, \dots, \ell\}$ , par  $L_{\ell,m}(u) := (-1)^m (1-u^2)^{m/2} \frac{d^m}{du^m} L_\ell(u)$ . Les fonctions  $L_{\ell,m}$  vérifient, comme d'ailleurs les polynômes de Legendre, une formule de récurrence permettant leur calcul.

On montre enfin qu'une base orthonormée de  $\mathbb{H}_\ell$  est formée par les fonction  $(\mathcal{Y}_{\ell,m})_{-\ell \leq m \leq \ell}$  où

$$\begin{cases} \mathcal{Y}_{\ell,m}(\theta, \varphi) := \sqrt{\frac{4\pi}{2\ell+1} \frac{(\ell-m)!}{(\ell+m)!}} L_{\ell,m}(\cos \theta) e^{im\varphi}, & m \geq 0 \\ \mathcal{Y}_{\ell,-m} := \mathcal{Y}_{\ell,m}^* & \text{(N.B. il est parfois ajouté un facteur } (-1)^m \text{ à la définition de } \mathcal{Y}_{\ell,-m}) \end{cases} \quad (2.3)$$

La suite complète  $(\mathcal{Y}_{\ell,m})_{\ell \in \mathbb{N}, -\ell \leq m \leq \ell}$  constitue donc une base orthonormée de  $\mathbb{L}^2(\mathbb{S}^2)$  d'après le théorème 2.1.

**Remarque 2.1.** (i) Il est clair à partir de (2.3) que si  $f$  est à valeurs réelles alors  $a_{\ell,-m} = a_{\ell,m}^*$ .

(ii) Pour les fonctions à valeurs réelles, on peut alternativement utiliser la base des harmoniques sphériques réelles, constituée, pour chaque fréquence  $\ell \in \mathbb{N}$ , des  $2\ell+1$  fonctions  $\mathcal{Y}_{\ell,0}$ ,  $\sqrt{2}\text{Re}(\mathcal{Y}_{\ell,m})$ ,  $\sqrt{2}\text{Im}(\mathcal{Y}_{\ell,m})$ ,  $m \in \{1, \dots, \ell\}$ .

(iii) Les fonctions définies par les équations (2.3) sont les plus utilisées en pratique. Dans toute la suite, elles peuvent néanmoins être remplacées par n'importe quelle autre base orthonormée de  $\mathbb{H}_\ell$ .

## Construction à partir de la théorie des groupes

D'autres constructions des harmoniques sphériques sont possibles. Il est parfois utile d'utiliser par exemple celle à partir des représentations du groupe  $SO(3)$  (voir par exemple [44]).

On peut repérer toute rotation  $\rho \in SO(3)$  par ses angles d'Euler  $(\phi, \theta, \psi) \in [0, 2\pi[ \times [0, \pi[ \times [0, 2\pi[$  car toute rotation s'écrit de manière unique  $u(\phi)a(\theta)u(\psi)$  où

$$u(\psi) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad a(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}.$$

On dispose d'une base orthonormée (de vecteurs propres de l'opérateur de Laplace-Beltrami sur  $SO(3)$ ) de  $\mathbb{L}^2(SO(3), d\rho)$  où  $d\rho$  est la mesure de Haar. Cette base est constituée des fonctions

$\rho = (\phi, \theta, \psi) \mapsto D_{\ell, m, m'}(\rho)$ ,  $\ell \in \mathbb{N}$ ,  $m, m' \in \{-\ell, \dots, \ell\}$  définies par

$$D_{\ell, m, m'}(\phi, \theta, \psi) := \frac{i^{m-m'}}{2^\ell} \sqrt{\frac{(\ell - m')!}{(\ell - m)!(\ell + m')!(\ell + m)!}} \times e^{-i(m\phi + m'\psi)} (\sin \theta)^{m'-m} (1 + \cos \theta)^m \left[ \frac{d^{\ell+m'}}{dx^{\ell+m'}} (x-1)^{\ell+m} (x+1)^{\ell-m} \right]_{x=\cos \theta}. \quad (2.4)$$

Il est alors possible de définir la transformation de Fourier sur  $SO(3)$  : pour  $g \in SO(3)$ ,

$$\begin{aligned} A_{\ell, m, m'}^g &:= \int_{SO(3)} g(\rho) D_{\ell, m, m'}(\rho) d\rho \\ g &= \sum_{\ell=0}^{\infty} (2\ell + 1) \sum_{m, m'=-\ell}^{\ell} A_{\ell, m, m'}^g D_{\ell, m, m'}^* \quad (\text{dans } \mathbb{L}^2(SO(3), d\rho)) \end{aligned}$$

Or on dispose d'une représentation unitaire de  $SO(3)$  dans  $\mathbb{L}^2(\mathbb{S}^2)$  :

$$\rho \in SO(3) \mapsto (f \in \mathbb{L}^2(\mathbb{S}^2) \mapsto f \circ (\rho^{-1}) \in \mathbb{L}^2(\mathbb{S}^2)) \quad (2.5)$$

En identifiant  $\mathbb{S}^2$  à  $SO(3)/SO(2)$ , la base  $\mathcal{Y}_{\ell, m}$  des harmoniques sphériques n'est autre que

$$\mathcal{Y}_{\ell, m}(\theta, \phi) = (4\pi)^{-1/2} (2\ell + 1)^{1/2} (-1)^m D_{\ell, m, 0}^*(\phi, \theta, \psi)$$

qui ne dépend pas de  $\psi$  (car  $m' = 0$  dans (2.4)).

Cela permet tout d'abord d'expliciter le point 1 du théorème 2.1 sur la base usuelle des harmoniques sphériques.

**Proposition 2.1.** *Soit  $\ell \in \mathbb{N}$ . Le sous espace engendré par les fonctions  $\mathcal{Y}_{\ell, m}$ ,  $-\ell \leq m \leq \ell$  est stable par rotation. Plus précisément, les sous-espaces harmoniques  $\mathbb{H}_\ell$  sont les composantes irréductibles de la représentation définie par l'équation (2.5), et pour toute rotation  $\rho \in SO(3)$ ,*

$$\mathcal{Y}_{\ell, m} \circ \rho = \sum_{m'=-\ell}^{\ell} (-1)^{m'} D_{\ell, m, m'}(\rho^{-1}) \mathcal{Y}_{\ell, m'}.$$

De plus, il est possible de définir la convolution entre une fonction  $f \in \mathbb{L}^2(\mathbb{S}^2)$  et une fonction  $g \in \mathbb{L}^2(SO(3), d\rho)$  et d'obtenir une généralisation du théorème 2.3. Soit en effet

$$g * f(\xi) := \int_{SO(3)} f(\rho^{-1}\xi) g(\rho) d\rho.$$

**Proposition 2.2.** *Pour toutes fonctions  $f \in \mathbb{L}^2(\mathbb{S}^2)$  et  $g \in \mathbb{L}^2(SO(3), d\rho)$ ,  $a_{\ell, m}^{g*f} = \sum_{m'=-\ell}^{\ell} A_{\ell, m, m'}^g a_{\ell, m'}^f$ , ou encore, sous forme matricielle,  $a_\ell^{g*f} = A_\ell^g a_\ell^f$ .*

## 2.2 Champs aléatoires sur la sphère

### Champs aléatoires et distributions directionnelles

Les études statistiques sur la sphère sont nombreuses et variées. Elles peuvent être regroupées en deux grandes catégories, selon que les variables aléatoires considérées sont à valeur dans  $\mathbb{S}^2$  ou indexées par



$\mathbb{S}^2$ . Dans le premier cas, on parle de statistiques directionnelles (ou plus précisément de statistiques sphériques). Dans le second, on parle de champs aléatoires (ou de processus) sphériques. Une distinction similaire existe en dimension inférieure, avec les processus stochastiques périodiques et l'étude de données directionnelles circulaires. Bien que fondamentalement différentes, ces deux catégories de problèmes peuvent bénéficier des mêmes techniques d'analyse sur la sphère. Un exemple récent est l'utilisation des needlets (qui seront définies dans la section 2.3) par Baldi, Kerkycharian, Marinucci et Picard à la fois dans le contexte de champs aléatoires sur la sphère et le cercle [151, 152, 149] et dans le contexte de statistiques directionnelles [150].

Malgré quelques travaux précurseurs dans les années 1960-70 [181, 163, 177, 178], il semble que l'étude des processus sur la sphère a été principalement, jusqu'à ces dernières années, le fait de physiciens (en météorologie, astrophysique et géophysique notamment). Cette étude sera discutée dans la suite de ce chapitre. Quant aux statistiques sphériques, elles ont fait l'objet de recherches depuis relativement longtemps. Des revues peuvent être trouvées dans les livres [172, chap. 8 et 9], [158], [180, chap. 10], [173, chap. 9 et 10].

Parmi la variété de problèmes de statistiques sphériques, l'estimation non paramétrique de densité est probablement celui qui est le plus susceptible de profiter de techniques similaires à l'étude des processus, puisqu'il s'agit d'estimer des fonctions de la sphère. Le problème non bruité [160, 148, 179, 162] est modélisé par l'observation de  $X_1, \dots, X_n$  i.i.d. de densité  $f$  inconnue (par rapport à  $d\xi$ ). Les vitesses minimax d'estimation de  $f$ , dans les boules de Besov sur la sphère, pour les pertes  $\mathbb{L}^p$ ,  $1 \leq p < \infty$ , et en norme sup, ainsi que la réalisation de ces vitesses par le seuillage des coefficients d'ondelettes estimés de  $f$  sont étudiées dans [150].

Dans le problème du bruit blanc, on observe cette fois  $dX(\xi) = f(\xi)d\xi + \epsilon dW(\xi)$  où  $\epsilon \in \mathbb{R}_+^*$  et  $dW$  est le bruit blanc gaussien sur  $\mathbb{S}^2$ . Klemelä [167] a calculé des vitesses minimax asymptotiques (quand  $\epsilon \rightarrow 0$ ), dans les boules de Hölder pour la perte en norme sup, et dans les boules de Sobolev pour la perte  $\mathbb{L}^2$  (voir aussi [168] pour des généralisations).

Toujours dans le cadre de l'estimation de densité, une autre façon de modéliser le bruit des observations est de considérer le modèle de convolution par un opérateur aléatoire :  $Z_1, \dots, Z_n$  i.i.d. selon  $Z \sim \epsilon X$  où  $X$  est de densité  $f$  sur  $\mathbb{S}^2$  toujours à estimer et  $\epsilon$  est un élément aléatoire de  $SO(3)$ . Dans ce problème, initié sur la sphère dans [161], Kim et Koo [166] ont calculé les vitesses minimax. Ces derniers [169] ont aussi généralisé récemment le problème du bruit blanc en un problème inverse :  $dX(\xi) = \Lambda f(\xi)d\xi + \epsilon dW(\xi)$ , où cette fois l'opérateur  $\Lambda$  est déterministe.

## Champs aléatoires stationnaires sur la sphère

Revenons au contexte qui nous intéresse directement pour la suite de ce travail. Rappelons ici quelques résultats sur l'analyse harmonique des processus aléatoires sphériques, dont les analogues dans le cas euclidien sont bien connus. Dorénavant, seule une famille de variables aléatoires indexées par  $\mathbb{S}^2$  et à valeurs réelles est considérée, c'est-à-dire un processus de la forme  $(\Omega, \mathcal{A}, \mathbf{P}, X(\xi), \xi \in \mathbb{S}^2)$  où  $(\Omega, \mathcal{A}, \mathbf{P})$  est un espace probabilisé et  $X$  prend ses valeurs dans  $\mathbb{R}$ . On suppose de plus ce processus centré et continu en moyenne quadratique. Autrement dit,

**Hypothèse 2.1.** *On suppose que pour tout  $\xi \in \mathbb{S}^2$ ,  $\mathbf{E} X(\xi) = 0$  et  $\mathbf{Var} X(\xi) < \infty$ , et que la fonction  $X : \mathbb{S}^2 \rightarrow \mathbb{L}^2(\Omega, \mathbf{P})$  est continue.*

En particulier (par compacité de  $\mathbb{S}^2$ )  $X$  est intégrable en moyenne quadratique (i.e. ses sommes de

Riemann convergent dans  $\mathbb{L}^2(\Omega, \mathbf{P})$ ). Appliquant les résultats de la section 2.1, on dispose alors de la décomposition en harmoniques sphériques (ou de Fourier) de  $X$ , c'est-à-dire de variables aléatoires

$$a_{\ell,m}^X := \int_{\mathbb{S}^2} X(\xi) \mathcal{Y}_{\ell,m}^*(\xi) d\xi$$

et

$$X_\ell(\xi) := \sum_{m=-\ell}^{\ell} a_{\ell,m}^X \mathcal{Y}_{\ell,m}(\xi),$$

et du résultat suivant [181].

**Proposition 2.3.** *Sous l'hypothèse 2.1, les variables aléatoires  $a_{\ell,m}^X$  et  $X_\ell(\xi)$ ,  $\ell \in \mathbb{N}$ ,  $m \in \{-\ell, \dots, \ell\}$ ,  $\xi \in \mathbb{S}^2$ , sont bien définies, et de plus  $X = \sum_{\ell=0}^{\infty} X_\ell$ , au sens où  $\mathbf{E} \int_{\mathbb{S}^2} \left( X(\xi) - \sum_{\ell=0}^L X_\ell(\xi) \right)^2 d\xi \xrightarrow{L \rightarrow \infty} 0$*

L'adaptation à la sphère des notions de stationnarité conduit aux deux définitions suivantes (qui sont équivalentes dans le cas où le processus est gaussien).

**Définition 2.5.** (i) *Le processus  $X$  est dit stationnaire<sup>1</sup> (au second ordre) s'il existe  $\gamma : [0, \pi] \rightarrow \mathbb{R}$ , appelée fonction de corrélation angulaire, telle que pour tous  $\xi, \xi' \in \mathbb{S}^2$ ,*

$$\mathbf{E} X(\xi) X(\xi') = \gamma(d(\xi, \xi')).$$

*Autrement dit, la corrélation entre deux directions ne dépend que de l'angle entre ces directions.*

(ii) *Le processus  $X$  est dit strictement stationnaire si pour tous  $N \in \mathbb{N}$ ,  $(\xi_1, \dots, \xi_N) \in (\mathbb{S}^2)^N$ , et pour toute rotation  $\rho \in SO(3)$ , les vecteurs  $(X(\xi_1), \dots, X(\xi_N))$  et  $(X \circ \rho(\xi_1), \dots, X \circ \rho(\xi_N))$  ont même loi.*

**Hypothèse 2.2.** *On suppose le processus  $X$  stationnaire (au second ordre).*

Comme dans le cas de séries temporelles, la stationnarité dans le domaine direct a comme contrepartie dans le domaine de Fourier la décorrélation des coefficients et l'identité de leur variance à une fréquence  $\ell$  donnée.

**Proposition 2.4.** *Sous les hypothèses 2.1 et 2.2, il existe une suite réelle positive  $C_\ell^X$ ,  $\ell \in \mathbb{N}$ , appelée spectre de puissance angulaire telle que  $\forall \ell, \ell' \in \mathbb{N}$ ,  $\forall m \in \{-\ell, \dots, \ell\}$ ,  $\forall m' \in \{-\ell', \dots, \ell'\}$ ,*

$$\mathbf{E} \operatorname{Re}(a_{\ell,m}^X) \operatorname{Re}(a_{\ell',m'}^X) = \delta_{\ell,\ell'} \delta_{m,m'} \left( \delta_{m,0} C_\ell^X + (1 - \delta_{m,0}) \frac{C_\ell^X}{2} \right),$$

$$\mathbf{E} \operatorname{Im}(a_{\ell,m}^X) \operatorname{Im}(a_{\ell',m'}^X) = \delta_{\ell,\ell'} \delta_{m,m'} (1 - \delta_{m,0}) \frac{C_\ell^X}{2},$$

$$\mathbf{E} \operatorname{Re}(a_{\ell,m}^X) \operatorname{Im}(a_{\ell',m'}^X) = 0.$$

La proposition 2.4, conséquence par exemple d'un résultat plus général sur les processus homogènes et isotropes dans  $\mathbb{R}^3$  démontré par Yaglom [181], a aussi été prouvée par Baldi et Marinucci [153] sans l'hypothèse de continuité en moyenne quadratique mais en supposant la stricte stationnarité.

<sup>1</sup>Certains auteurs disent aussi "isotrope" et d'autres "homogène".

**Corollaire 2.2.** *Sous les hypothèses 2.1 et 2.2,  $\forall \ell, \ell' \in \mathbb{N}, \forall m \in \{-\ell, \dots, \ell\}, \forall m' \in \{-\ell', \dots, \ell'\},$*

$$\mathbf{E} a_{\ell,m}^X (a_{\ell',m'}^X)^* = \delta_{\ell,\ell'} \delta_{m,m'} C_\ell^X. \quad (2.6)$$

*En particulier,*

$$\forall \ell, \ell' \in \mathbb{N}, \forall \xi, \xi' \in \mathbb{S}^2, \mathbf{E} X_\ell(\xi) X_{\ell'}(\xi') = \delta_{\ell,\ell'} C_\ell^X L_\ell(\xi \cdot \xi')$$

*et*

$$\gamma(\theta) = \sum_{\ell=0}^{\infty} C_\ell^X L_\ell(\cos \theta). \quad (2.7)$$

A nouveau, la formule (2.7) est l'analogie du lien entre le spectre et la fonction d'autocorrélation des séries temporelles stationnaires, où la série de Fourier est ici remplacée par une série “de Legendre”. Une autre conséquence de la proposition 2.4 concerne les processus gaussiens, pour lesquels les coefficients de Fourier  $a_{\ell,m}$  sont également gaussiens (complexes).

**Corollaire 2.3.** *Sous les hypothèses 2.1 et 2.2, si  $X$  est gaussien, les variables  $a_{\ell,m}^X$ ,  $\ell \in \mathbb{N}$ ,  $m \in \{0, \dots, \ell\}$  sont deux à deux indépendantes.*

Rappelons cependant que malgré l'équation (2.6) les variables  $a_{\ell,m}^X$  et  $a_{\ell,-m}^X$  ne sont pas indépendantes car elles sont conjuguées l'une de l'autre (Remarque 2.1 (i)).

**Remarque 2.2.** *L'Hypothèse 2.1 impose la sommabilité de  $(\ell C_\ell^X)_{\ell \in \mathbb{N}}$ .*

## Exemples d'applications

Comme cela a été mentionné en introduction de ce chapitre, on rencontre des champs aléatoires sphériques dans de très nombreuses applications. L'étude du Fond diffus cosmologique a été présentée au chapitre 1. Les quelques autres exemples qui suivent ne sont pas destinés à donner une vue d'ensemble de ces domaines, mais à en illustrer la diversité. Ils sont choisis parmi les travaux récents, certains seront réutilisés à la fin de ce chapitre à propos des ondelettes sphériques.

En géophysique, les variations spatiales du champ magnétique terrestres sont étudiées, ainsi que leur évolution dans le temps. Le champ magnétique est mesuré en permanence par plusieurs dizaines de stations d'observation au sol, réparties de façon très inhomogène sur la surface du globe ; des mesures sont également réalisées à partir de satellites. Ces dernières fournissent des données globales et plus homogènes, mais aussi moins bien résolues et moins précises. En effet, le champ magnétique créé par la lithosphère est plus faible, à l'altitude des satellites, comparé aux autres sources – dont la séparation semble d'ailleurs être un enjeu. Voir par exemple [189]. Les articles de Holschneider, Chambodut et Mandeau [184] et de Thébault [189] discutent de techniques d'analyse permettant de reconstituer, à partir des données, le champ magnétique, globalement et localement, aux grandes et aux petites échelles angulaires (jusqu'à des fréquences  $\ell$  typiquement de l'ordre de quelques dizaines).

En météorologie et en climatologie, un problème en partie similaire se rencontre, le champ magnétique étant remplacé par le champ de température (ou de pression) à la surface du globe. C'est d'ailleurs également l'un des domaines ayant motivé en premier l'étude de champs aléatoires sur la sphère [156, 178]. La modélisation météorologique nécessite souvent un paramètre supplémentaire temporel et considère des processus  $X(\xi, t)$ ,  $\xi \in \mathbb{S}^2$ ,  $t \in T \subset \mathbb{R}$ . L'étude de ces processus fut initiée, avec  $T = \mathbb{Z}$ , par Jones [163] dans le cas i.i.d. et Roy [178] dans le cas stationnaire ; Jun et Stein [164, 165] considèrent

des covariances spatio-temporelles plus générales. Néanmoins, le modèle d'un champ  $X(\xi)$  est aussi utilisé pour des champs de température instantanée ou au contraire moyenne (sur une ou plusieurs années), comme par exemple dans les articles de Oh et Li [186, 185], dans lesquels par ailleurs le point commun avec les questions liées au champ magnétique ressort clairement. En effet, les températures sont mesurées par un réseau de stations elles aussi réparties de façon inhomogène sur la Terre.

L'expression *données éparpillées* ("scattered data") décrit cette grande inhomogénéité des points d'observation. C'est un trait commun entre les données géophysiques et météorologiques, qui marque une différence avec le cas du Fond Diffus cosmologique. Ce dernier est généralement observé sur des zones plus ou moins grandes, mais avec un échantillonnage présentant plus de régularité.

En imagerie médicale, mais aussi dans l'industrie et en vision artificielle, on modélise la forme d'objets tridimensionnels en projetant leur surface sur une sphère. Par exemple Yu et al. [191] considèrent la forme du cerveau, mesurée par Imagerie par Résonance Magnétique (IRM), dont l'évolution dans le temps, chez un individu, peut être le symptôme de pathologies. Ces auteurs relèvent l'importance de savoir représenter les caractéristiques globales et locales de cette surface de façon à pouvoir aligner les images de différentes mesures et détecter les variations malgré le bruit des mesures.

En astrophysique enfin, le *CMB* n'est pas le seul phénomène à être modélisé par un champ aléatoire sphérique. C'est aussi le cas de la répartition de matière – plus précisément de celle des galaxies. La répartition spatiale des galaxies est tridimensionnelle, mais sa composante sphérique est plus accessible que sa composante radiale. C'est d'ailleurs dans le cadre de l'étude statistique de cette répartition que Peebles [119] a introduit l'estimateur  $\widehat{C}^{PCL}$  défini au Chapitre 3 Eq. (3.9). L'article de Pietrobon, Balbi et Marinucci [121] étudie la corrélation entre le champ modélisant le Fond diffus et celui modélisant la répartition de galaxies.

## 2.3 Ondelettes sphériques

### Motivations

L'utilité, tant théorique que pratique, des ondelettes en traitement du signal et en statistique n'est plus à démontrer. Classiquement, l'analyse en ondelette peut être construite sur  $\mathbb{R}$  ou  $\mathbb{R}^n$ , par translation et dilatation de fonctions mères, permettant d'analyser de façon localisée temporellement (ou spatialement) et fréquemment. Le terme "ondelette" sera ici entendu au sens large de fonction jouissant de propriétés de localisation dans ces deux domaines. Sur la sphère, l'intérêt potentiel de telles fonctions ressort des propriétés du Fond diffus cosmologique introduites au Chapitre 1 et qui seront précisées au Chapitre 3 (données manquantes, mélange de processus dont certains ne sont pas stationnaires, importance de l'estimation spectrale, ...) et des enjeux dans les autres applications mentionnées dans la section précédente.

Malheureusement, la généralisation des ondelettes "euclidiennes" à la sphère n'est pas immédiate. Notamment le problème d'une définition opérationnelle de la dilatation se pose sur ce compact qu'est la sphère, comme il se pose d'ailleurs, par exemple, sur un intervalle de longueur finie (mais avec la difficulté supplémentaire qu'on ne peut pas "périodiser" des fonctions de la sphère). Une première possibilité est de projeter stéréographiquement la sphère (moins un point) sur le plan et de réaliser la dilatation dans le plan. Ainsi, la dilatation d'un facteur  $r$  autour du pôle Nord  $N(\theta = 0)$  est

$$\xi \mapsto S^{-1}(rS(\xi))$$

où  $S$  est la bijection

$$S : (\theta, \varphi) \in \mathbb{S}^2 \setminus N \mapsto 2 \tan \frac{\theta}{2} e^{i\varphi} \in \mathbb{C}. \quad (2.8)$$

Cette transformation sera appelée *dilatation stéréographique*. Elle conduit, quand  $r \rightarrow \infty$ , à une accumulation au pôle Sud, ce qui est inévitable lorsque l'on cherche à dilater d'un facteur important sur  $\mathbb{S}^2$ . Par contre, l'intérêt est que dans la limite  $r \rightarrow 0$  on retrouve immédiatement la dilatation euclidienne.

Une autre possibilité de définition de la dilatation découle de l'observation suivante : dans le cas euclidien, à toute dilatation dans le domaine direct correspond (à un facteur multiplicatif près) une dilatation inverse dans le domaine de Fourier. Adapté à la sphère, cela abouti, pour une fonction  $f \in \mathbb{L}^2(\mathbb{S}^2)$  dont les coefficients de Fourier sont de la forme  $a_{\ell,m}^f = \mathbf{b}(\ell)$ , à la dilatation  $\mathbf{b}(\ell) \mapsto \mathbf{b}(r^{-1}\ell)$ .

Ces deux possibilités, et d'autres encore, ont été explorées. Dans ce travail, l'intérêt se portera essentiellement sur la seconde.

Quant aux propriétés de localisation, elles peuvent être définies de diverses façons. Ce qui est parfois appelé localisation, au sens le plus strict, consiste pour une fonction à avoir un support inclus dans une boule pour la distance  $d$  (appelée également calotte sphérique) de rayon inférieur à  $\pi$  (et que l'on peut donc rendre arbitrairement petit par dilatation stéréographique). On peut considérer l'analogue dans l'espace de Fourier : une fonction dont le support des coefficients de Fourier est fini est dite *band-limited*. Comme dans le cas euclidien, une fonction non nulle ne peut pas être à la fois band-limited et strictement localisée, car un polynôme s'annulant sur une calotte sphérique de rayon non nul s'annule sur toute la sphère.

Un critère de localisation moins contraignant consiste à étudier la “variance” d'une fonction. Cette dernière permet d'obtenir un principe d'indétermination sur la sphère analogue à celui portant sur les fonctions de la variable réelle. Il est dû à Narcowich et Ward [34].

**Définition 2.6.** Soit  $f \in \mathbb{L}^2(\mathbb{S}^2)$ ,  $\|f\|_2 = 1$ . Le centre de gravité de la mesure  $|f(\xi)|^2 d\xi$  est le vecteur

$$\mu(f) := \int_{\mathbb{S}^2} \xi |f(\xi)|^2 d\xi.$$

La variance correspondante est

$$\sigma_\xi^2(f) := \int_{\mathbb{S}^2} \|\xi - \mu(f)\|^2 |f(\xi)|^2 d\xi.$$

Définissons également la quantité

$$\sigma_\ell^2(f) := \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \ell(\ell+1) a_{\ell,m}^2.$$

**Remarque 2.3.** La définition de  $\sigma_\ell^2(f)$ , qui peut sembler arbitraire, est en fait une expression, dans le domaine de Fourier, de la variance du rotationnel en  $f$  (voir [34, 20]).

Le centre de gravité appartient à la boule unité de  $\mathbb{R}^3$ , et sa norme est d'autant plus grande que  $f$  est localisée autour du point  $\|\mu(f)\|^{-1}\mu(f)$ .

**Lemme 2.4.** Soit  $f \in \mathbb{L}^2(\mathbb{S}^2)$ ,  $\|f\|_2 = 1$ . Alors  $\sigma_\xi^2(f) = 1 - \|\mu(f)\|^2$ .

Le théorème suivant quantifie le fait qu'une fonction ne peut pas être parfaitement localisée simultanément en espace et en fréquence spatiale (le fait que 1 soit la meilleure constante pour la partie droite de l'inégalité (2.9) est prouvé dans [20]).

**Théorème 2.4.** *Soit  $f \in \mathbb{L}^2(\mathbb{S}^2)$  de norme 1 et de classe  $C^2$ . Alors*

$$\sigma_{\xi}(f) \sigma_{\ell}(f) \geq \|\mu(f)\|. \quad (2.9)$$

Le troisième critère de localisation envisagé est la concentration en norme, c'est-à-dire, étant donné une région  $R \subset \mathbb{S}^2$ , la quantité

$$C_R(f) := \frac{\|f \mathbf{1}_R\|_2^2}{\|f\|_2^2} \quad (2.10)$$

Wieczorek, Simons et Dahlen [190, 41] ont montré que si l'on se restreint à des fonctions band-limited,  $f \in \bigoplus_{\ell=0}^L \mathbb{H}_{\ell}$ , alors l'optimisation du critère  $C_R$  se réduit à un problème de valeurs propres d'une matrice symétrique définie positive. Une suite orthogonale de fonctions optimalement concentrées dans  $R$  s'en déduit. Le problème dual, de concentration dans le domaine de Fourier parmi les fonctions de support inclus dans  $R$ , s'en déduit par simple troncature.

Enfin, la localisation peut être mesurée par la vitesse de décroissance de la queue d'une fonction loin de son maximum, c'est-à-dire trouver une vitesse  $v$  mesurant la localisation de  $f$  au point  $\xi_0$  par

$$|f(\xi_0)|^{-1} |f(\xi)| \leq v(d(\xi, \xi_0)).$$

## Premiers travaux

Le développement d'ondelettes sur la sphère, bien que récent, a déjà donné lieu à une littérature trop vaste pour qu'une présentation exhaustive des travaux existants puisse tenir en quelques pages. Du moins, une présentation historique, et espérons assez complète, des premiers travaux dans cette direction peut aider à en saisir les enjeux.

La construction d'ondelettes sphériques a surgi sous la forme d'une vingtaine d'articles entre 1995 et 1998, reflétant des tentatives dans des directions très différentes et parfois complémentaires. Ces travaux étaient en partie motivés par les applications en géophysique, et le groupe de Géomathématique de l'Université de Kaiserslautern (Freedon et al.) a d'ailleurs notablement contribué à ces premiers développements.

La première tentative de définition d'ondelettes sur la sphère semble être celle de Dahlke, Dahmen, Schmitt et Weinreich [7]. Il s'agit d'un produit tensoriel d'ondelettes euclidiennes en  $\theta \in [0, \pi]$  et périodiques en  $\varphi \in \mathbb{R}/2\pi\mathbb{Z}$ . Plus précisément, une Analyse Multirésolution (AMR) sur la sphère est engendrée par

$$\Phi_{j,(k_1,k_2)}(\theta, \varphi) := \Phi_{j,k_1}^1(\theta) \Phi_{j,k_2}^2(\varphi)$$

où les  $\Phi_{j,k_1}^1$  génèrent une AMR sur  $\mathbb{L}^2([0, \pi])$  et  $\Phi_{j,k_2}^2(\varphi) := \sum_{k \in \mathbb{Z}} \eta(2^j(\varphi - k) - k_1)$  avec  $\eta$  générant une AMR sur  $\mathbb{L}^2(\mathbb{R})$ . Les conditions aux bords sont résolues de façon que  $\Phi_{j,(k_1,k_2)}$  soit de classe  $C^1$  sur  $\mathbb{S}^2$ . La même idée (avec des fonctions  $\Phi^1$  et  $\Phi^2$  différentes) fut poursuivie par Potts et Tasche [38] (qui passèrent peu après [37] à une définition plus intrinsèque de la transformation en ondelettes, grâce à l'utilisation d'une dilatation dans le domaine de Fourier). Les singularités aux pôles des coordonnées  $(\theta, \varphi)$  posent cependant des difficultés qui n'ont pas pu être totalement surmontées.

Schöder et Sweldens [39] construisent des ondelettes par le procédé de “lifting scheme” sur des partitions hiérarchiques de la sphère, ce qui permet une construction intrinsèque évitant les singularités inhérentes à tout système de coordonnées sphériques, comme celui des angles  $(\theta, \phi)$ . L’approche est algorithmique, et vise essentiellement la synthèse numérique d’images (“computer graphics”). L’algorithme permet de décomposer et reconstruire une fonction à partir de ses valeurs locales, selon un schéma “biorthogonal” (au sens où les filtres d’analyse et de synthèse diffèrent). Göttelmann [23] reprend cette méthode et la complète mathématiquement. En définissant des splines sur des ensembles de points d’échantillonnage définis pour l’occasion, il obtient des vitesses d’approximation dans les espaces de Sobolev pour l’interpolation par ces splines et en déduit une AMR. Il montre qu’on peut obtenir pour cette AMR des frames qui soient localisées, c’est-à-dire à décroissance polynomiale :

$$|\psi_{j,k}(\xi)| \leq c \left(1 + 2^j d(\xi, \xi_{j,k})\right)^{-1},$$

ou même strictement localisées. L’inconvénient est qu’il n’y a en contrepartie aucun contrôle de la localisation dans le domaine dual (c’est-à-dire des coefficients de Fourier de  $\psi_{j,k}$ ).

L’approche de Torrèsani [45, 46]<sup>2</sup>, motivée par la théorie des groupes, consiste en une transformée de Fourier à fenêtre sphérique, indexée par  $SO(3) \times \mathbb{R}$  (ou  $\mathbb{S}^2 \times \mathbb{R}$  dans le cas d’une fenêtre radiale). La modulation de la fenêtre est faite dans le plan tangent et une formule d’analyse-reconstruction est montrée. Holschneider [26] utilise lui aussi le plan tangent à la sphère, mais pour définir une sorte de dilatation ad hoc sur la sphère (qui n’est pas la dilatation stéréographique) au moyen de transformations de Fourier sur  $\mathbb{R}^2$  et sur les cercles. Cela rend possible la définition d’une transformation continue (biorthogonale) en ondelettes jouissant des propriétés de conservation d’énergie, de reconstruction exacte et de réduction de la dilatation sphérique considérée en la dilatation stéréographique dans la limite des petites échelles : il existe une fonction  $g \in L^2(\mathbb{R}^2)$  telle que

$$r^2 g_r(S^{-1}(rx)) \xrightarrow{r \rightarrow 0} g(x)$$

où  $g_r \in L^2(\mathbb{S}^2)$  définit la dilatation d’un facteur  $r$  sur  $\mathbb{S}^2$  et  $S$  est définie en (2.8). Une troisième approche utilisant le plan tangent pour définir une transformation sphérique continue en ondelettes (qui est ensuite discrétisée) est due à Dahlke et Maass [8].

L’objectif d’Antoine et Vandergheynst [5, 4] est, d’une part, la production d’une analyse continue par rotation et dilatation d’une ondelette mère, en définissant simplement et géométriquement la dilatation sphérique, et d’autre part, l’obtention de la transformée euclidienne usuelle comme limite lorsque  $\mathbb{S}^2$  peut être localement identifiée à son plan tangent. Ils démontrent que la dilatation stéréographique est en fait bien adaptée à cette tâche. Des arguments de théorie des groupes justifient ce point de vue : il permet de définir les ondelettes comme états cohérents de  $\mathbb{S}^2$  (c’est-à-dire images dans  $L^2(\mathbb{S}^2)$  d’une représentation unitaire) associés à l’action des rotations de  $SO(3)$ , et de  $\mathbb{R}_+^*$  par la dilatation stéréographique. Les représentations en question, considérées sur des sphères de rayon  $R$ , se contractent vers la représentation usuelle dans  $L^2(\mathbb{R}^2)$  du groupe des similitudes de  $\mathbb{R}^2$ , quand  $R \rightarrow \infty$ .

Malgré leur intérêt conceptuel, aucune des approches précédentes n’aboutit à une localisation quantifiée des ondelettes dans les domaines direct *et* de Fourier.

L’article de Narcowich et Ward [34] développe une méthode adaptée à l’analyse de données éparpillées, comme celles rencontrées en géophysique et en météorologie (voir la fin de la section 2.2). Une AMR (indépendante du choix de coordonnées) est construite à partir

<sup>2</sup>Mentionnons, selon le souhait de l’auteur, que cette construction peut être obtenue comme cas particulier d’une construction antérieure plus générale due à de Bièvre [10].

- d’une “fonction de base sphérique” (FBS)  $\Lambda : [-1, 1] \rightarrow \mathbb{R}$ , telle que le noyau associé  $(\xi, \xi') \mapsto \Lambda(\xi \cdot \xi')$  soit défini positif sur  $\mathbb{S}^2$ ,
- d’une suite emboîtée d’ensemble de points d’échantillonnage, dense dans la sphère. Cela équivaut, quitte à réindexer les points, à supposer une suite  $(\xi_k)_{k \in \mathbb{N}}$  de points telle que  $\overline{\bigcup_{k \in \mathbb{N}} \xi_k} = \mathbb{S}^2$ .

L’AMR est engendrée par un nombre croissant des premiers points  $\xi_k$  sur lesquels sont centrées les fonctions radiales associées à  $\Lambda$  (selon la définition 2.2). Les auteurs discutent de la possibilité de construction d’une base orthonormée d’ondelettes associée à cette AMR. Par ailleurs, pour évaluer les qualités de ces ondelettes, ils introduisent et démontrent le principe d’incertitude sur la sphère énoncé dans le théorème 2.4. Dans le cas particulier de la FBS  $\Lambda^{AP(p)} := \sum_{\ell=0}^{\infty} p^\ell L_\ell$ ,  $0 < p < 1$ , la localisation spatiale de l’ondelette associée à l’AMR est calculée au sens de ce principe d’incertitude :  $\mu(\psi_k) = \frac{2p}{1+p^2} \xi_k$ . Autrement dit, le centre de masse de l’ondelette est aligné avec le point d’échantillonnage considéré, et sa variance peut être rendue arbitrairement faible (ce qui se paye, intuitivement, par un nombre de points d’échantillonnage d’autant plus grand pour décrire correctement des fonctions sur toute la sphère).

Cette approche ressemble à celle des needlets dont il est question plus loin. Mais cette AMR est construite à partir des seules rotations (celles qui centrent les fonctions radiales  $\psi_k$  sur les points  $\xi_k$ ), sans faire intervenir de dilatation. Il n’y a donc pas de phénomène de localisation s’améliorant avec l’échelle.

Freedman et al., dans une série d’articles [17, 19, 20, 16, 18, 183], adoptent également le point de vue de la convolution par des fonctions radiales, mais contrairement à [34] une famille  $\Lambda_r$  de FBS est utilisée et la dilatation est définie par cette famille ; de plus, dans le cas de la transformation discrète, les points à chaque échelle sont supposés relativement bien répartis et non pas éparpillés. Le point de départ est l’approximation de l’identité, c’est-à-dire une famille  $\Lambda_r \in \mathbb{L}^2([-1, 1])$  telle que

$$\Lambda_r * f \xrightarrow{r \rightarrow 0} f$$

pour toute  $f \in \mathbb{L}^2(\mathbb{S}^2)$ . Par exemple,  $\Lambda_r = \Lambda^{AP(e^{-r})}$  définie ci-dessus convient (le noyau associé est appelé noyau d’Abel-Poisson, d’où la notation  $\Lambda^{AP}$ ).

Dans un premier temps [19], une transformation continue en ondelettes est construite à partir des fonctions d’échelles  $\phi_{r,\eta}(\xi) := \Phi_r(\eta \cdot \xi)$  où  $\Phi_r := \Lambda_r^{1/2}$  et des ondelettes  $\psi_{r,\eta} := \Psi_r(\eta \cdot \xi)$  où  $\Psi_r \in \mathbb{L}^2([-1, 1])$  est telle que  $\left((\Phi_r)_\ell^\wedge\right)^2 = \left(\int_R^\infty \left((\Psi_r)_\ell^\wedge\right)^2 dr\right)^{1/2}$ .<sup>3</sup> Sous réserve que la condition d’admissibilité

$$\int_{\mathbb{R}_+} \left((\Psi_r)_\ell^\wedge\right)^2 dr = 1 \quad (2.11)$$

soit satisfaite<sup>3</sup>, il est possible de définir une AMR continue analogue au cas euclidien. La dilatation est donc définie relativement à l’indice de la famille de noyaux et la rotation consiste à centrer sur l’ensemble des points de  $\mathbb{S}^2$  les fonctions radiales associées aux noyaux.

Des versions semi-discrètes sont également proposées, selon plusieurs méthodes dont les paquets d’ondelettes, avec des familles dénombrables de noyaux  $\Psi_j$  et  $\Phi_j$ ,  $j \in \mathbb{N}$ . La condition d’admissibilité (2.11) devient dans ce cas

$$\sum_{j=-\infty}^{\infty} \left((\Psi_j)_\ell^\wedge\right)^2 = 1. \quad (2.12)$$

Un schéma biorthogonal est possible en remplaçant  $\left((\Psi_r)_\ell^\wedge\right)^2$  par  $(\Psi_j)_\ell^\wedge (\widetilde{\Psi}_j)_\ell^\wedge$  dans (2.12). Enfin, en



imposant des échelles dyadiques et les propriétés  $\Phi_J * \Phi_J = \Phi_{J-1}$  et  $\Psi_j * \widetilde{\Psi}_j = \Psi_{j-1}$ , on obtient un schéma récursif pyramidal de décomposition et reconstruction à partir de l'échelle la plus fine.

La discrétisation de l'échelle spatiale au moyen de formules de quadrature est présentée dans [20]. La transformée discrète obtenue n'utilise des fonctions à analyser que leurs valeurs sur les points d'échantillonnage  $(\xi_1, \dots, \xi_N)$ . Elle ne fournit en général qu'une formule de reconstruction approchée, dont la convergence, quand  $N \rightarrow \infty$ , est établie sous certaines conditions portant sur la répartition des points  $\xi_k$ , et pour des fonctions appartenant à des espaces de Sobolev sur la sphère. Dans le cas d'une ondelette à support fréquentiel fini, une transformée discrète et exacte est possible.

Cette idée est développée par Freedon et Schreiner [18]. Dans cet article, la dilatation est systématiquement définie dans l'espace de Fourier<sup>4</sup>, au moyen d'une fonction continue par morceaux  $\mathbf{b} : \mathbb{R}_+ \rightarrow [0, 1]$  qui est dilatée dyadiquement et discrétisée en

$$b_{j,\ell} := \mathbf{b}(2^{-j}\ell). \quad (2.13)$$

Les nombres  $b_{j,\ell}$ ,  $\ell \in \mathbb{N}$ , définissent les coefficients de Legendre du noyau  $\Psi_j$  :

$$\Psi_j := \sum_{\ell=0}^{\infty} b_{j,\ell} L_{\ell}. \quad (2.14)$$

La condition d'admissibilité (2.12) est vérifiée si par exemple  $\mathbf{b}(\cdot) = \sqrt{\mathbf{a}(\cdot/2) - \mathbf{a}(\cdot)}$  avec  $\mathbf{a}(0) = 1$  et  $\mathbf{a}$  continue en 0. Si de plus  $\mathbf{b}$  est à support compact :  $\text{supp}(\mathbf{b}) \subset [0, C]$ ,  $\Psi$  est dite *band-limited*. Dans ce cas, une formule discrète exacte d'analyse-synthèse découle de formules de cubature. Les ondelettes  $\psi_{j,k}$  correspondant à ce schéma ont depuis été appelées *needlets* par Narcowich, Petrushev et Ward [33, 32] qui ont démontré leur localisation quasi-exponentielle (polynomiale à tout ordre) pour une fonction  $\mathbf{a}$  régulière et des formules de cubature bien choisies, et le fait que ces ondelettes constituent une frame ajustée ("tight frame").

## Les needlets : propriétés de localisation et propriétés statistiques

Dans la dernière décennie, les différentes constructions d'ondelettes sphériques ont connu de nombreux développements théoriques, des réalisations numériques et de premières fructueuses applications. Par exemple, les ondelettes continues obtenues par dilatation stéréographique ont été généralisées en des ondelettes directionnelles [3, 50]. Leur utilisation pour l'analyse du Fond diffus cosmologique est devenue courante, notamment pour des tests de stationnarité et de gaussianité [65, 28, par exemple]. Starck et al. [134] ont développé des curvlets et ridgelets sur la sphère en s'appuyant sur une transformée dans le domaine de Fourier et sur le schéma hiérarchique de pixelisation HEALPix [84]. Les exemples d'applications mentionnés à la fin de la section 2.2 utilisent différentes formes d'ondelettes. Dernier exemple, Geller et Mayeli [21, 175] ont étendu la construction des needlets à des variétés riemanniennes compactes et, en modifiant légèrement la définition, trouvé leur formule explicite dans le domaine direct (au prix, par contraste avec ce qui va suivre, de l'abandon de leur caractère "band-limited").

Ce qui fait défaut aux différentes sortes d'ondelettes énumérées ci-dessus est l'absence d'un contrôle simultané sur leur localisation fréquentielle (e.g. band-limited) et spatiale. C'est pourquoi l'obtention

<sup>3</sup>L'intégration sur  $\mathbb{R}_+^*$  en  $r$  est généralement écrite avec une mesure  $w(r)dr$  où, par exemple,  $w(r) = r^{-1}$  par analogie avec les ondelettes euclidiennes. Ce choix n'est pas nécessaire [20, p. 12] et ici le poids  $w(r)$  est conventionnellement inclus dans  $(\Psi_r)_{\ell}^{\wedge}$ .

<sup>4</sup>Signalons que cette idée est déjà mentionnée par Torrèsani [45, chap. VII.VII]

par Narcowich, Petrushev et Ward [33, 32] d'une décroissance quasi-exponentielle (polynomiale à tout ordre) pour des needlets band-limited a été particulièrement utile dans le travail présenté ici. Cette double décroissance assure de bonnes propriétés d'approximation et d'un point de vue statistique la décorrélation asymptotique des coefficients de needlet d'un champ stationnaire. Concluons ce chapitre en rappelant ces résultats.

Partons d'une fonction  $\mathbf{a} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  décroissante et d'un réel  $\mathcal{B} > 1$ . Posons

$$\mathbf{b}(\cdot) = \sqrt{\mathbf{a}(\cdot/\mathcal{B}) - \mathbf{a}(\cdot)} . \quad (2.15)$$

Originellement,  $\mathcal{B}$  était choisi égal à 2. Une valeur arbitraire de  $\mathcal{B}$  a été introduite par la suite pour plus de souplesse dans l'analyse du CMB.

Définissons  $\Psi_j$  par l'équation (2.14) où cette fois

$$b_{j,\ell} := \mathbf{b}(\mathcal{B}^{-j}\ell) . \quad (2.16)$$

Une nouveauté est l'utilisation de formules de cubature avec des poids positifs, obtenues dans [29, 33].

**Proposition 2.5.** *Pour tout  $\ell^{max} \in \mathbb{N}$ , il existe  $N \in \mathbb{N}$ , des points  $(\xi_1, \dots, \xi_N) \in (\mathbb{S}^2)^N$  et des poids positifs  $(\lambda_1, \dots, \lambda_N) \in \mathbb{R}_+^N$  tels que*

$$\forall f \in \bigoplus_{\ell=0}^{\ell^{max}} \mathbb{H}_\ell, \quad \int_{\mathbb{S}^2} f(\xi) d\xi = \sum_{k=1}^N \lambda_k f(\xi_k)$$

et  $c_0^{-1} (\ell^{max})^2 \leq N \leq c_0 (\ell^{max})^2$  où  $c_0 \in \mathbb{R}_+^*$  est une constante indépendante de  $\ell^{max}$ .

Selon cette proposition, donnons-nous à chaque échelle  $j$  des couples  $(\xi_{j,k}, \lambda_{j,k})$ ,  $k \in \{1, \dots, N_j\}$ , fournissant des formules de cubature pour  $\ell^{max}$  valant  $\mathcal{B}^{2j}$ . Les needlets et les coefficients de needlet d'une fonction  $f \in \mathbb{L}^2(\mathbb{S}^2)$  sont respectivement définis par

$$\psi_{j,k} : \xi \in \mathbb{S}^2 \mapsto \sqrt{\lambda_{j,k}} \Psi_j(\xi \cdot \xi_k) \in \mathbb{R} \quad \text{et} \quad \beta_{j,k} := \int_{\mathbb{S}^2} f(\xi) \psi_{j,k}(\xi) d\xi . \quad (2.17)$$

Parmi les propriétés de ces fonctions, retenons celles résumées dans le théorème suivant.

**Hypothèse 2.3.** *La fonction  $\mathbf{a}$  (Eq. 2.15) est de classe  $C^M$ ,  $2 \leq M \leq \infty$ , et telle que  $\mathbf{a}(x) = 1$  si  $x \leq \mathcal{B}^{-1}$  et  $\mathbf{a}(x) = 0$  si  $x > 1$ .*

Dans ce cas, les équations (2.15)–(2.16) définissent une décomposition de type Littlewood-Paley.

**Théorème 2.5.** *Sous l'hypothèse 2.3*

1. *Les fonctions  $\psi_{j,k}$ ,  $j \in \mathbb{N}$ ,  $k \in \{1, \dots, N_j\}$  constituent une frame ajustée de  $\mathbb{L}^2(\mathbb{S}^2)$  :  $\forall f \in \mathbb{L}^2(\mathbb{S}^2)$ ,*

$$f(\xi) = \sum_{j=-\infty}^{\infty} \sum_{k=1}^{N_j} \beta_{j,k} \psi_{j,k}(\xi) , \quad (2.18)$$

$$\|f\|_2 = \sum_{j=-\infty}^{\infty} \sum_{k=1}^{N_j} \beta_{j,k}^2 . \quad (2.19)$$

2. *Pour tout  $m < M$ , il existe une constante  $C_m$  telle que*

$$\forall j \in \mathbb{N}, k \in \{1, \dots, N_j\}, \quad \forall \xi \in \mathbb{S}^2, \quad |\psi_{j,k}(\xi)| \leq \frac{C_m \mathcal{B}^j}{(1 + \mathcal{B}^j d(\xi, \xi_k))^m} . \quad (2.20)$$

L'utilisation de cette dernière propriété aboutit à un résultat immédiat mais remarquable : les coefficients de needlets d'un champ stationnaire sont deux à deux asymptotiquement décorrélés (sauf en des points distants d'un ordre de grandeur de  $\mathcal{B}^{-j}$  ou moins). Soit en effet  $X$  un champ aléatoire stationnaire, de spectre de puissance  $(C_\ell)_{\ell \in \mathbb{N}}$ .

**Hypothèse 2.4.** *Il existe  $\alpha > 2$  et une suite de fonctions  $\mathbf{g}_j : [\mathcal{B}^{-1}, \mathcal{B}] \rightarrow \mathbb{R}$ ,  $j \in \mathbb{N}$ , telles que*

$$C_\ell = \ell^{-\alpha} \mathbf{g}_j(\mathcal{B}^{-j} \ell)$$

*pour tout  $\ell \in [\mathcal{B}^{j-1}, \mathcal{B}^{j+1}]$ . Il existe aussi des nombres positifs  $c_0, \dots, c_M$  tels que pour tout  $j \in \mathbb{N}$ ,  $c_0^{-1} \leq \mathbf{g}_j \leq c_0$  et pour tout  $m \leq M$ ,  $\sup_{\mathcal{B}^{-1} \leq u \leq \mathcal{B}} \left| \frac{d^m}{du^m} \mathbf{g}_j(u) \right| \leq c_m$ .*

**Corollaire 2.4.** *Soit  $X$  un champ stationnaire sur la sphère et  $\beta_{j,k}$  ses coefficients de needlets. Sous les hypothèses 2.3 et 2.4,*

$$|\mathbf{Cor}(\beta_{j,k}, \beta_{j,k'})| \leq \frac{C_m}{(1 + \mathcal{B}^j d(\xi_{j,k}, \xi_{j,k'}))^m}.$$

Cette propriété est utilisée par Baldi, Kerkycharian, Marinucci et Picard [151] pour établir un théorème de limite centrale pour des fonctionnelles des coefficients  $\beta_{j,k}$ . Elle servira aussi dans le chapitre 7 à établir la consistance d'estimateurs du spectre de puissance.

## Chapitre 3

# Enjeux statistiques liés à l'étude du Fond diffus cosmologique

### Sommaire

3.1	Modélisation du Fond diffus et des observations . . . . .	48
	Une réalisation unique, cas idéal . . . . .	48
	Modélisation des observations . . . . .	49
3.2	Estimation du spectre de puissance angulaire . . . . .	50
	Retour sur la motivation physique . . . . .	50
	Données manquantes et bruitées . . . . .	51
	Bruit de niveau inconnu et inter-spectres . . . . .	53
	Remarques . . . . .	54
3.3	Extraction du Fond diffus (séparation de source) . . . . .	55
	Modèle . . . . .	55
	Principe d'estimation . . . . .	56
	Discussion . . . . .	58
3.4	Autres enjeux statistiques . . . . .	58
	Projection sphérique de données temporelles corrélées . . . . .	58
	Estimation des paramètres cosmologiques . . . . .	59
	Séparation de sources . . . . .	59
	Test des hypothèses du modèle (stationnarité, gaussianité) . . . . .	60

La trace des fluctuations du *CMB* sur la surface de dernière diffusion est généralement modélisée par la réalisation d'un champ sphérique aléatoire (gaussien, le plus souvent) : pour un espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$ , on considère le processus aléatoire indexé par la sphère  $X(\xi)$ ,  $\xi \in \mathbb{S}^2$ , que l'on suppose centré et continu en moyenne quadratique (Hypothèse 2.1).

Suivant le principe cosmologique (voir page 11), le champ  $X$  est supposé stationnaire<sup>1</sup> (la stationnarité au second ordre, Hypothèse 2.2, nous suffira dans la suite). Dans ce cas, la structure de covariance du

<sup>1</sup>Nous utilisons le terme “stationnaire”, et non “isotrope”, car le principe cosmologique postule une isotropie en un sens légèrement différent (non statistique) : il est question dans notre cas de l'isotropie (la stationnarité)... des “anisotropies” (selon le vocabulaire physique).

champ (i.e. sa fonction d'autocorrélation) est entièrement caractérisée par son *spectre de puissance angulaire* (Proposition 2.4, Corollaire 2.2)

Le spectre de puissance est le résumé le plus important de la loi du processus, c'est à partir de ses valeurs estimées que la plupart des inférences statistiques sont réalisées, par exemple l'estimation des paramètres cosmologiques. Dans le cas où le processus est supposé gaussien et donc caractérisé par ses propriétés du second ordre, le spectre de puissance détermine totalement la loi de  $X$ . Or les modèles physiques standards d'inflation génèrent justement un champ gaussien de fluctuations.

L'estimation du spectre de puissance, problème standard dans le cas de processus temporels (ou "euclidiens"), est rendue difficile ici, outre la géométrie originale du problème, par les modèles d'observation imposés par les expériences physiques (voir la Section 1.3). Nous tenons compte ci-dessous de ces conditions d'observation dans notre modélisation.

Nous présentons de façon séparée deux problèmes statistiques : celui de l'estimation du spectre de puissance à partir d'observations bruitées et partiellement masquées ; puis celui de la séparation de champs sphériques indépendants à partir de l'observation de plusieurs mélanges de ces champs. D'autres questions statistiques liées à l'étude du *CMB* sont mentionnées à la fin de ce chapitre.

### 3.1 Modélisation du Fond diffus et des observations

#### Une réalisation unique, cas idéal

Avant de modéliser les observations de  $X$ , insistons sur le fait que, même dans le cas idéal de l'observation complète et non bruitée de  $X$ , une seule réalisation de ce processus nous est accessible<sup>2</sup>.

Des asymptotiques sont cependant envisageables. Une première est de considérer les hautes fréquences spatiales :  $\ell \rightarrow \infty$ .

**Définition 3.1.** On appelle spectre empirique la suite

$$\widetilde{C}_\ell(X) := (2\ell + 1)^{-1} \sum_{m=-\ell}^{\ell} |a_{\ell,m}^X|^2.$$

**Proposition 3.1.**  $\widetilde{C}_\ell(X)$  est un estimateur non biaisé de  $C_\ell^X$ . De plus, si  $X$  est gaussien, alors  $\widetilde{C}_\ell(X)$  est l'estimateur du maximum de vraisemblance et, si  $C_\ell^X > 0$ ,

$$\frac{(2\ell + 1)}{C_\ell^X} \widetilde{C}_\ell(X) \sim \chi_{2\ell+1}^2.$$

En particulier,

$$\frac{\mathbf{E} \left( \widetilde{C}_\ell(X) - C_\ell^X \right)^2}{(C_\ell^X)^2} = \frac{2}{2\ell + 1} \quad (3.1)$$

L'équation (3.1) révèle la propriété remarquable suivante, liée à la géométrie de la sphère (à travers le fait que  $\dim(\mathbb{H}_\ell) = 2\ell + 1$ , Théorème 2.1) : malgré l'unicité de la réalisation du processus, l'estimation

---

<sup>2</sup>En effet le découplage ne s'est réalisé qu'une fois, et la durée qui nous en sépare est fixée, à l'échelle humaine.

du spectre par le spectre empirique converge aux hautes fréquences, au sens où :

$$\frac{\mathbf{E} \left( \widetilde{C}_\ell(X) - C_\ell^X \right)^2}{\left( C_\ell^X \right)^2} \xrightarrow{\ell \rightarrow \infty} 0 \quad . \quad (3.2)$$

Une deuxième asymptotique, qui n'est pas sans lien avec la première comme on va le voir, consiste à limiter ou contrôler les nuisances (bruit, couverture incomplète et inhomogène, nombre fini de valeurs observées, etc.) dans l'observation du processus d'intérêt. Nous précisons ci-dessous la modélisation de ces nuisances.

### Modélisation des observations

Les “nuisances” dans l'observation de  $X(\xi)$ . sont multiples. Nous énumérons ci-dessous les principales.

1. Tout d'abord, l'observation n'étant pas infiniment résolue, on ne peut observer que

$$B * X(\xi) := \int_{\mathbb{S}^2} X(\xi') B(\xi, \xi') d\xi' \quad (3.3)$$

pour une certaine réponse impulsionnelle  $B$ .

Cela correspond à un modèle inverse (de convolution).

Si  $B$  est à symétrie circulaire, c'est-à-dire  $B(\xi, \xi') = \widetilde{B}(\xi \cdot \xi')$  avec  $\widetilde{B} : [-1, 1] \rightarrow \mathbb{R}$ , alors (3.3) s'exprime dans le domaine de Fourier par  $a_{\ell, m}^{B * X} = \widetilde{B}_\ell^\wedge a_{\ell, m}^X$  et  $C_\ell^{B * X} = \left( \widetilde{B}_\ell^\wedge \right)^2 C_\ell^X$ .

2. Comme des avant-plans émettent un rayonnement qui se mélange au *CMB*, les fluctuations de température équivalente mesurées de la lumière reçue peuvent s'écrire en chaque direction  $B * (X + S)(\xi)$ , où  $S(\xi)$  est la contribution de diverses sources astrophysiques.
3. Seules des observations spatialement discrètes, définies sur des points (ou pixels)  $\xi_p$ ,  $1 \leq p \leq N_{pix}$  sont accessibles.
4. Pour tenir compte de la couverture incomplète du ciel ou de l'application d'un masque sur certaines zones, les observations sont multipliées par une fonction  $W_p = W(\xi_p)$ , qui peut être choisie de la forme  $W(\xi) = \mathbf{I}(\xi \in R)$  où  $R \subset \mathbb{S}^2$  est la région observée, ou bien d'une forme plus générale<sup>3</sup>.
5. Le bruit instrumental hétéroscédastique entachant l'observation de  $X(\xi_p)$  s'écrit  $\sigma_p U_p$  où  $\sigma_p \in \mathbb{R}_+$  et  $U_p \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ .

Les points 1, 3 et 5 ci-dessus sont directement reliés à la fréquence spatiale maximale “observable”. En effet les fréquences  $\ell$  pour lesquelles  $B_\ell = 0$  sont perdues et celles pour lesquelles  $B_\ell$  est très petit par rapport au niveau de bruit (point 5) ne peuvent être retrouvées. De toute façon, le calcul numérique des coefficients de Fourier à partir d'un nombre fini  $N_{pix}$  de points n'est possible que jusqu'à une valeur maximale de  $\ell$ , qui dépend de  $N_{pix}$  (et même de l'ensemble des points  $\xi_1, \dots, \xi_{N_{pix}}$ ). Pour une expérience donnée, il ne saurait donc être question d'asymptotique à haute fréquence.

<sup>3</sup>En pratique, les masques, originellement à valeur dans  $\{0, 1\}$ , sont souvent apodisés pour limiter les effets de bord (“spectral leakage”).

Cependant, le point 1 reflète le niveau technologique atteint dans la résolution angulaire d'une expérience. Quant au point 3, il dépend du temps et de la précision d'observation (car il y a un compromis à faire, dans la stratégie d'observation et de fabrication de cartes, entre le nombre de points mesurés et le nombre de fois que chaque point est mesuré, dont dépend le niveau de bruit effectif) ; il dépend également de la limite informatique à la taille des cartes sur lesquelles les calculs sont possibles.

En résumé, les expériences qui se succèdent, en laissant  $B(\xi, \xi')$  s'approcher d'un Dirac,  $N_{pix}$  tendre vers l'infini, et  $\sigma$  tendre vers 0, donnent accès à des fréquences spatiales maximales  $\ell_{max}$  toujours plus grandes<sup>4</sup>.

Tenant compte du fait que les expériences mesurent toutes le même Fond diffus, avec plusieurs détecteurs, dans plusieurs bandes de fréquences lumineuses, tout modèle réaliste des observations doit donc finalement s'écrire, au minimum,

$$Y_{e,p,v,d} = W_{e,p} \left( B_{e,v} * (X + S_v)(\xi_{e,p}) + \sigma_{e,p,v,d} U_{e,p,v,d} \right) \quad (3.4)$$

où  $e \in \{1, \dots, N_{exp}\}$  indexe l'expérience,  $p \in \{1, \dots, N_{pix}(e)\}$  numérote les points d'observation,  $v \in \{1, \dots, N_{freq}(e)\}$  et  $d \in \{1, \dots, N_{det}(e, v)\}$  sont respectivement pour la fréquence lumineuse et le détecteur, et les  $U_{e,p,v,d}$  sont centrées, de variance 1 et indépendantes de  $X$  et  $S$ .

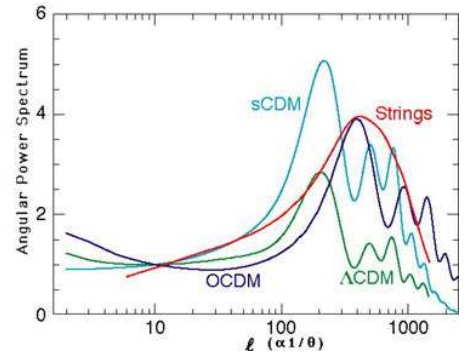
Il est bien sûr possible d'étudier des modèles d'observations simplifiés, ne reprenant que certains des paramètres de (3.4). C'est ce que nous faisons dans la suite.

## 3.2 Estimation du spectre de puissance angulaire

### Retour sur la motivation physique

Maintenant que nous avons défini le spectre de puissance, revenons sur l'intérêt physique de son estimation, ébauché dans le Chapitre 1. Ce sont les progrès dans son estimation qui ont permis les avancées les plus importantes en cosmologie. La raison en est que le spectre est une fonction des paramètres cosmologiques qui, même si elle n'est pas connue analytiquement, peut être calculée numériquement. De nombreuses méthodes, comme CMBFAST [130] ou CAMB [101] résolvent numériquement les équations déterminant le spectre avec précision et rapidité (permettant leur utilisation dans des procédures de Monte Carlo).

La connaissance du spectre permet tout d'abord de départager différentes classes de modèles cosmologiques (voir la figure 3.2 – le modèle favorisé par les



F . 3.1 – Exemples de spectres de puissance  $C_\ell^X$ ,  $2 \leq \ell \leq 2000$ , pour différents modèles cosmologiques<sup>5</sup>. [image : Smoot]

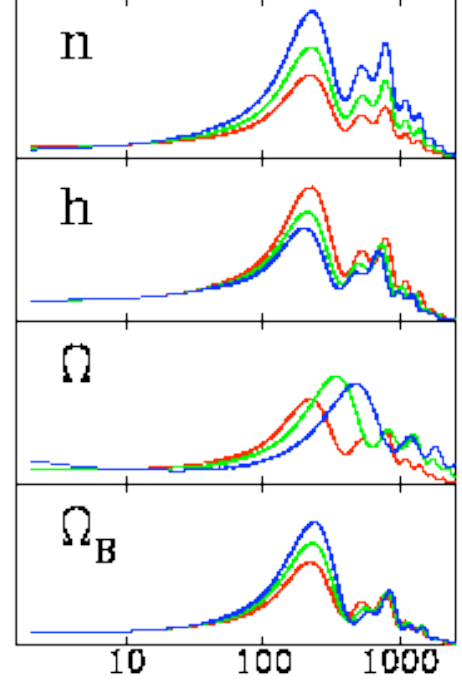
<sup>4</sup> Pour reprendre quelques exemples déjà évoqués :  $\ell_{max} = 2$  dans les années 1970 ;  $\ell_{max} = 30$  pour COBE (1992) ;  $\ell_{max} = 700$  pour ARCHEOPS (2002) ;  $\ell_{max} = 1000$  pour WMAP (2003) ;  $\ell_{max} = 2000$  attendu pour PLANCK ;  $\ell_{max} = 3500$  attendu pour QUIET.

<sup>5</sup> Suivant la tradition, ce sont les fonctions  $\ell \mapsto \frac{\ell(\ell+1)}{2\pi} C_\ell$  qui sont représentées.

observations actuelles étant le modèle  $\Lambda$ CDM<sup>6</sup>). Ensuite, pour un modèle donné, le spectre permet d'en estimer les paramètres.

La figure 3.2 illustre la dépendance du spectre de puissance en fonction des paramètres cosmologiques du modèle  $\Lambda$ CDM. La caractéristique principale de ces spectres est la présence de pics, qui représentent, comme expliqué dans la section 1.2, les échelles spatiales dont les oscillations ont connu un nombre entier de demi-périodes. Le premier pic correspond à la fréquence spatiale fondamentale des oscillations (c'est-à-dire à l'échelle pour laquelle une seule compression s'est produite jusqu'au découplage), et les pics suivants aux harmoniques (pour lesquels il y a eu une succession de compressions – décompressions).

Intuitivement, la hauteur du premier pic est déterminée, entre autres, par la densité de matière, une plus grande densité augmentant l'amplitude de la compression. Autre exemple, l'évolution de la proportion de matière dans le temps peut rendre les oscillations asymétriques : si par exemple la compression devient plus importante que la décompression, les pics impairs gagnent en hauteur par rapport aux pics pairs.



F . 3.2 – Exemples de spectres de puissance  $C_\ell^X$ ,  $2 \leq \ell \leq 2000$ , pour différentes valeurs de certains paramètres cosmologiques<sup>5</sup>. [image : White]

### Données manquantes et bruitées

Comme on le voit, l'estimation du spectre de puissance peut apporter des informations précieuses. Dans le cas idéal de l'observation de  $X(\xi)$ ,  $\xi \in \mathbb{S}^2$ , sous l'hypothèse gaussienne, la Proposition 3.1 justifie l'utilisation du simple spectre empirique, basé sur les coefficients de Fourier  $a_{\ell,m}^X$ . Mais les données manquantes ou masquées, ainsi que la présence de bruit compliquent l'estimation. Principalement, les données manquantes biaisent tous les coefficients de Fourier et les corrélient, ce qui rend inutilisable en l'état le spectre empirique  $\tilde{C}_\ell(Y)$ .

Considérons le modèle bruité et masqué suivant (le masque  $W$  et le niveau de bruit  $\sigma$  sont supposés connus) :

$$Y_p = W_p (X(\xi_p) + \sigma_p U_p) \quad , \quad 1 \leq p \leq N_{pix} \quad (3.5)$$

Deux classes d'estimateurs sont utilisés en pratique pour estimer le spectre  $C^X := (C_\ell^X)_{\ell \in \mathbb{N}}$  dans le modèle (3.5) : le maximum de vraisemblance dans le domaine direct et la déconvolution du spectre empirique (avec au préalable une pondération des données). Notons  $Y := (Y_1, \dots, Y_{N_{pix}})$ .

**Proposition 3.2.** *Dans le modèle (3.5), si  $X$  est gaussien,  $U_p \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$  et  $W_p = \mathbf{I}(p \in P)$ , la*

<sup>6</sup> $\Lambda$ CDM pour “ $\Lambda$  and Cold Dark Matter”, c'est-à-dire avec de la matière noire froide et une constante cosmologique suffisante pour rendre l'espace plat.



vraisemblance du spectre est

$$\mathcal{L}(C^X|Y) = \left( (2\pi)^{\tilde{N}_{pix}} |K| \right)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \tilde{Y}^T K^{-1} \tilde{Y} \right) ,$$

où  $K_{p,q} := \sum_{\ell=0}^{\infty} C_{\ell}^X L_{\ell}(\xi_p \cdot \xi_q) + \sigma_p^2 \delta_{p,q}$ ,  $p, q \in P$ ,  $\tilde{N}_{pix} := \#P$  et  $\tilde{Y}$  est le vecteur formé des éléments de  $Y$  dont les indices sont dans  $P$  (sans perte de généralité, on peut supposer avoir ordonné les points de sorte que  $P = \{1, \dots, \tilde{N}_{pix}\}$ ).

L'inconvénient pratique est que cette méthode nécessite l'inversion de  $K$ , de taille  $\tilde{N}_{pix} \times \tilde{N}_{pix}$ . Ceci n'est donc réalisable en pratique que jusqu'à des tailles  $\tilde{N}_{pix}$  de quelques milliers, ce qui ne permet que l'estimation des fréquences spatiales inférieures à quelques dizaines. L'estimation du spectre à partir des données de COBE a pu être réalisée de cette façon, et il est toujours possible, pour des cartes de grande taille, de calculer l'estimateur  $\widehat{C}^{DML}(Y)$  maximisant la vraisemblance de cartes dégradées (c'est-à-dire projetées sur un moindre nombre de pixels). Les possibilités d'implémentation de telles méthodes, ou de méthodes approchant le maximum de vraisemblance ont été largement débattues (voir par exemple [137, 58]), mais leur complexité reste de l'ordre de  $\tilde{N}_{pix}^3$ . Elles ne permettent donc pas l'estimation de  $C_{\ell}^X$  au-delà de  $\ell \simeq 60$  (avec les calculateurs actuels).

Une autre approche consiste à débiaiser et déconvoluer, dans le domaine de Fourier, le spectre empirique. Cette approche, appelée dans la littérature astrophysique le “*pseudo- $C_{\ell}$* ”, est efficace en pratique mais n'a reçu que peu de traitement statistique rigoureux.

**Hypothèse 3.1.** –  $X$  est band-limited (en moyenne quadratique), avec  $\forall \ell > \ell^{max}$ ,  $a_{\ell,m}^X = 0$   
–  $(\xi_p, \lambda_p)_{p \in \{1, \dots, N_{pix}\}}$  assure une cubature à l'ordre  $2\ell^{max}$ .

**Définition 3.2.** Pour un  $N_{pix}$ -uplet  $(x_1, \dots, x_{N_{pix}})$ , on note

$$\begin{aligned} - a_{\ell,m}(x) &:= \sum_{p=1}^{N_{pix}} \lambda_p x_p \mathcal{Y}_{\ell,m}(\xi_p) \\ - \widetilde{C}_{\ell}(x) &:= (2\ell + 1)^{-1} \sum_{m=-\ell}^{\ell} |a_{\ell,m}(x)|^2 \end{aligned}$$

Remarquons que sous l'Hypothèse 3.1, si  $x_p = X(\xi_p)$ , alors  $a_{\ell,m}(x) = a_{\ell,m}^X$ .

**Proposition 3.3.** Sous l'Hypothèse 3.1, si  $Y = (Y_1, \dots, Y_{N_{pix}})$  suit le modèle (3.5), alors

$$\mathbf{E} \widetilde{C}_{\ell}(Y) = \sum_{\ell'=0}^{\ell^{max}} \mathcal{W}_{\ell,\ell'}(W) C_{\ell'}^X + N_{\ell}(W\sigma) \quad (3.6)$$

où

$$\mathcal{W}_{\ell,\ell'}(W) := \sum_{p=1}^{N_{pix}} \lambda_p W_p \sum_{m,m'=-\ell}^{\ell} \mathcal{Y}_{\ell,m}(\xi_p) \mathcal{Y}_{\ell,m'}(\xi_p) \quad (3.7)$$

et

$$N_{\ell}(V) := \mathbf{E} (2\ell + 1)^{-1} \sum_{m=-\ell}^{\ell} \left| \sum_{p=1}^{N_{pix}} \lambda_p V_p U_p \mathcal{Y}_{\ell,m}(\xi_p) \right|^2 . \quad (3.8)$$

La matrice  $\mathcal{W}(W)$  est dite de couplage. Dans le cas où les variables  $U_p$  sont i.i.d., le biais dû au bruit s'exprime plus simplement :

$$N_{\ell}(W\sigma) = (4\pi)^{-1} \sum_{p=1}^{N_{pix}} \lambda_p^2 W_p^2 \sigma_p^2 .$$

Compte tenu du fait que l'égalité (3.6) reste valable si les observations  $Y_p$  ont été préalablement pondérées par des poids  $V_p$  (c'est-à-dire en remplaçant  $Y$  par  $VY$  et  $W$  par  $VW$ ), on déduit de la proposition 3.3 une classe d'estimateurs non biaisés de  $C^X$  :

**Définition 3.3.** Soit, pour une pondération  $V$  et un masque  $W$  tels que l'inversion matricielle soit possible, l'estimateur

$$\widehat{C}_\ell^{PCL}(Y, V) := ([W(VW)]^{-1} \widetilde{C}(VY))_\ell - N_\ell(VW\sigma) \quad (3.9)$$

où  $W$  et  $N_\ell$  sont définis par les équations (3.7) et (3.8) respectivement (“PCL” est mis pour “pseudo- $C_\ell$ ”).

Des choix de  $V$  utilisés en pratique (par exemple pour estimer le spectre à partir des données WMAP [87]) sont  $V_p^{(1)} \equiv 1$  ;  $V_p^{(2)} = \sigma_p^{-2}$  ou encore  $V_p^{(3)} = \left( \sigma_p^2 + 4\pi \left( \sum_{k=1}^{N_{pix}} \lambda_k \sigma_k^{-2} \right)^{-1} \right)^{-1}$ .

Efstathiou [74] montre, dans le cas gaussien, que  $\widehat{C}_\ell^{PCL}(Y, V^{(1)})$  est asymptotiquement de variance minimale quand  $\sigma \rightarrow 0$  et que  $\widehat{C}_\ell^{PCL}(Y, V^{(2)})$  l'est asymptotiquement quand  $\sigma \rightarrow \infty$ , et il ajoute : “mais dans le régime intermédiaire il est difficile de trouver la façon de pondérer les données [... En effet] dans le domaine de Fourier,  $S$  [la matrice de covariance de  $X$ , en l'occurrence de  $a_{\ell,m}^X$ ] est creuse alors que  $N$  [celle du bruit] ne l'est pas ; à l'inverse dans le domaine direct on peut supposer que  $N$  est creuse mais  $S$  ne l'est pas”. L'auteur propose également une stratégie pour réaliser une transition entre  $\widehat{C}^{DML}(Y)$  et  $\widehat{C}_\ell^{PCL}(Y, V^{(1)})$ .

L'estimateur  $\widehat{C}_\ell^{PCL}(Y, V^{(3)})$  est une façon heuristique de réaliser une transition entre les deux régimes  $V^{(1)}$  et  $V^{(2)}$ . Dans notre cas,  $C_\ell^X \xrightarrow{\ell \rightarrow \infty} 0$  donc l'asymptotique de bruit faible est éventuellement valide pour les faibles fréquences  $\ell$  tandis que l'asymptotique de bruit fort l'est pour les grandes valeurs de  $\ell$ . En pratique, les fréquences limites auxquelles il vaut mieux passer d'un estimateur à l'autre sont estimées par des simulations de Monte Carlo. Par exemple, dans le traitement des données WMAP [87], les transitions entre les estimateurs  $\widehat{C}^{DML}(Y)$ ,  $\widehat{C}_\ell^{PCL}(Y, V^{(1)})$ ,  $\widehat{C}_\ell^{PCL}(Y, V^{(2)})$  et  $\widehat{C}_\ell^{PCL}(Y, V^{(3)})$  sont respectivement réalisées à  $\ell = 30$ ,  $\ell = 200$  et  $\ell = 450$ .

## Bruit de niveau inconnu et inter-spectres

Jusqu'ici les fonctions  $W$  et  $\sigma$  ont été supposées connues. Pour  $W$ , cela est réaliste car il s'agit d'un masque choisi. Par contre, le bruit n'est d'une part pas réellement i.i.d. et d'autre part pas nécessairement bien spécifié. Une façon de faire face à cette difficulté est d'utiliser, quand cela est possible, la redondance des observations. Pour cela, généralisons la définition 3.1 à l'*inter-spectre* empirique :

**Définition 3.4.** Soient deux champs sphériques  $X$  et  $X'$ . L'*inter-spectre empirique* entre  $X$  et  $X'$  est la suite  $(\widetilde{C}_\ell(X, X'))_{\ell \in \mathbb{N}}$  où

$$\widetilde{C}_\ell(X, X') := (2\ell + 1)^{-1} \sum_{m=-\ell}^{\ell} a_{\ell,m}^X (a_{\ell,m}^{X'})^* \quad (3.10)$$

Ainsi,  $\widetilde{C}_\ell(X) = \widetilde{C}_\ell(X, X)$ .

Considérons alors le modèle suivant, qui généralise celui de l'équation (3.5) :

$$Y_{p,d} = W_p(X(\xi_p) + \sigma_{p,d} U_{p,d}) \quad , \quad 1 \leq p \leq N_{pix} \quad , \quad 1 \leq d \leq N_{det} \quad (3.11)$$

où  $N_{det} > 1$  est le nombre de détecteurs. Définissons l'estimateur suivant qui, contrairement à celui de la Définition 3.3, ne nécessite pas la connaissance du niveau de bruit.

**Définition 3.5.** Définissons  $Y_d := (Y_{1,d}, \dots, Y_{N_{pix},d})$  et, pour une pondération  $V$  telle que l'inversion matricielle soit possible, l'estimateur

$$\widehat{C}_\ell^{CS}(Y_d, Y_{d'}, V) := \left( [\mathcal{W}(VW)]^{-1} \widetilde{C}(VY_d, VY_{d'}) \right)_\ell$$

(où “CS” est mis pour “cross-spectrum”).

**Proposition 3.4.** Dans le modèle (3.11), si les vecteurs  $U_d := (U_{1,d}, \dots, U_{N_{pix},d})$  sont indépendants de  $X$  et indépendants 2 à 2 (i.e. les bruits de deux détecteurs différents sont indépendants), alors  $\widehat{C}_\ell^{CS}(Y_d, Y_{d'}, V)$  est un estimateur non biaisé de  $C_\ell^X$ .

Dans le cas où, dans (3.11) il n'y a pas de masque ( $W \equiv 1$ ),  $X$  et  $U$  sont gaussiens et où le bruit est stationnaire, Polenta et al. [122] ont montré que la variance de l'estimateur

$$\frac{2}{N_{det}(N_{det} - 1)} \sum_{1 \leq d < d' \leq N_{det}} \widehat{C}_\ell^{CS}(Y_d, Y_{d'}, V^{(1)}) \quad (3.12)$$

est toujours supérieure à celle de l'estimateur

$$\frac{1}{N_{det}} \sum_{d=1}^{N_{det}} \widehat{C}_\ell^{PCL}(Y_d, V^{(1)}), \quad (3.13)$$

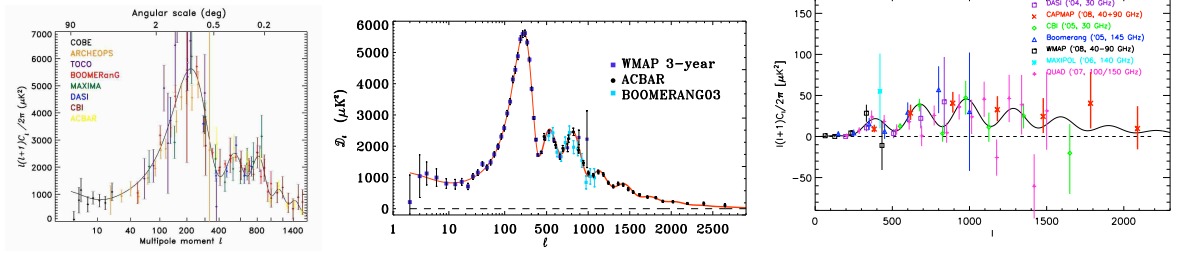
ce dernier supposant en revanche  $N_\ell$  connu. Les auteurs proposent un test de la spécification du bruit, basé sur la comparaison de ces deux estimateurs, de façon à décider entre l'utilisation de l'estimateur (3.13), plus efficace si le bruit est connu, et celle de l'estimateur (3.12), robuste.

## Remarques

Parmi les autres méthodes d'estimation, citons :

- Des méthodes bayésiennes explorant la loi jointe a posteriori du spectre par des simulations MCMC [77, 92].
- Des méthodes “multitaper” utilisant plusieurs fenêtres de pondérations de  $Y$  et moyennant les spectres empiriques correspondant. Dans le cas d'une observation masquée avec un bruit stationnaire, on peut [51, 182] ne pas chercher à déconvoluer le spectre empirique comme dans (3.9) mais à rendre les matrices  $\mathcal{W}$  aussi diagonales que possible en choisissant pour  $V$  l'équivalent sphérique des suites de Slepian (“discrete prolate spheroidal functions”). Les suites de spectres empiriques sont ensuite moyennées de façon à minimiser la variance de l'estimateur final.
- Une fois données les valeurs d'un vecteur  $\widehat{C}_\ell$ ,  $1 \leq \ell \leq \ell_{max}$ , estimant  $C_\ell^X$ , une régression non-paramétrique de  $C^X$  a été proposée [110, 159], selon le modèle  $\widehat{C}_\ell = f(\ell/\ell_{max}) + \varepsilon_\ell$ ,  $1 \leq \ell \leq \ell_{max}$ ,  $\varepsilon \sim \mathcal{N}(0, \Sigma)$ . Des ensembles de confiance asymptotiques pour des fonctionnelles de  $f$  (i.e. de  $C^X$ ) sont calculés pour  $\ell_{max} \rightarrow \infty$ .

Terminons cet aperçu des méthodes d'estimation spectrale par deux remarques. Premièrement, la plupart de ces méthodes reposent sur l'hypothèse de gaussianité de  $X$ , ce qui est l'hypothèse physique privilégiée actuellement. Cependant, si celle-ci n'était pas vérifiée, Baldi et Marinucci [153] ont montré que les coefficients  $a_{\ell,m}^X$  sont nécessairement non-indépendants.



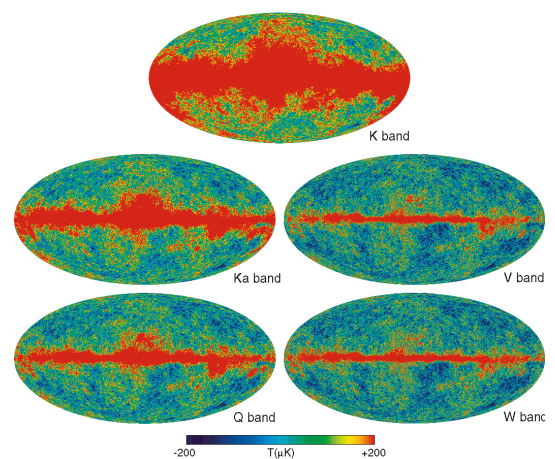
F . 3.3 – Exemples de résultats d’estimation spectrale publiés récemment. Les spectres estimés à partir de différentes expériences sont en couleurs différentes, avec des barres d’erreur indiquant les intervalles de confiance à 68%. Le spectre paramétrique résultant de l’estimation des paramètres cosmologiques à partir de ces différents spectres est ajouté en trait plein. La figure de droite correspond au spectre du mode dit  $E$  de la polarisation du  $CMB$ . [images, de gauche à droite : Hinshaw et al. 2003 [87] ; Reichardt et al. 2008 [124] ; Samtleben 2008]

Deuxièmement, la présentation faite ci-dessus de ces méthodes n’a fait intervenir que des versions très simplifiées du modèle d’observations (3.4). L’utilisation jointe de plusieurs observations, pourtant intuitivement pertinente, n’a été utilisée qu’à travers les inter-spectres (Définition 3.5). Bien sûr, l’information jointe de plusieurs expériences, via leurs spectres estimés respectifs, est utilisée pour estimer les paramètres cosmologiques, qui en retour fournissent un spectre estimé paramétrique. Mais il semble qu’aucune estimation de  $C^X$  basée sur plusieurs expériences (avec des niveaux de bruits, des beams et des masques différents) sans passer par les paramètres cosmologiques n’ait été présentée avant celle du Chapitre 6. De ce fait, les résultats de plusieurs expériences sont systématiquement présentés comme les exemples de la figure 3.3.

### 3.3 Extraction du Fond diffus (séparation de source)

#### Modèle

Revenons au modèle général (3.4). La contribution des avant-plans  $S_\nu$  dans les données, contrairement au bruit instrumental, n’est pas réduite par la multiplication des expériences et l’accroissement de la sensibilité des mesures. Au fil du temps et des progrès technologiques, la contribution des avant-plans, vue comme une contamination indésirable de la variable d’intérêt  $X$ , prend donc une importance croissante, et doit être traitée avec plus de précision car les résidus de cette contamination dans les cartes du  $CMB$  peuvent sérieusement biaiser toute interprétation cosmologique. Dans le paragraphe précédent, on a supposé une soustraction parfaite des avant-plans. Si tel n’est pas le cas, les spectres calculés seraient biaisés (positivement si les avant-plans sont supposés indépendants du Fond diffus). Notons que les techniques basées sur les inter-spectres sont impuissantes au regard des avant-



F . 3.4 – Cartes de température mesurées par l’expérience WMAP à cinq longueurs d’onde. [image : WMAP]

plans puisque ce sont les mêmes avant-plans qui sont vus par tous les détecteurs (à une fréquence lumineuse donnée).

La possibilité de séparer les avant-plans des fluctuations du Fond diffus vient de ce qu'ils n'ont pas, contrairement à ce dernier, un spectre lumineux dérivé de corps noir, et que l'on observe le ciel à plusieurs longueurs d'onde, indexées par  $\nu \in \{1, \dots, N_{freq}\}$  (voir l'annexe en fin de chapitre). La figure 3.4 illustre par exemple les cartes mesurées à cinq longueurs d'onde différentes par WMAP.

On considère le modèle contaminé et bruité suivant :

$$Y_{\nu,p} = X(\xi_p) + S_\nu(\xi_p) + \sigma_{\nu,p} U_{\nu,p}, \quad 1 \leq \nu \leq N_{freq}, \quad 1 \leq p \leq N_{pix} \quad (3.14)$$

où pour tout  $\nu$  les variables  $X$  et  $S_\nu$  sont supposées indépendantes, et où les variables de bruit  $U_{\nu,p}$  sont indépendantes entre elles et des précédentes. On cherche à retrouver  $X(\xi_p)$ , par une statistique  $\widehat{X}_p$  des observations des  $Y_{\nu,p}$ . Il s'agit donc d'un problème de séparation de source dans le cas particulier où on ne cherche à isoler qu'une seule source.

## Principe d'estimation

Nous cherchons un estimateur  $\widehat{X}_p$  comme combinaison linéaire des  $N_{freq}$  valeurs mesurées :

$$\widehat{X}_p(Y, w_p) := w_p^T Y_p \quad (3.15)$$

pour un certain vecteur de poids  $[w_{p,\nu}]_{1 \leq \nu \leq N_{freq}}$ .

Notons que dans le cas où toutes les variables en lettres capitales ci-dessus sont gaussiennes, l'espérance conditionnelle  $\mathbf{E} [X(\xi_p) | Y_p]$  est de cette forme avec  $(w_p^*)^T = R_{X_p Y_p} R_{Y_p}^{-1}$  où  $R_{Y_p} := \mathbf{E} Y_p Y_p^T$  et  $R_{X_p Y_p} := \mathbf{E} X(\xi_p) Y_p^T$ . Ce vecteur de poids vérifie en outre

$$w_p^* = \arg \min_w \mathbf{E} \left( \widehat{X}_p(Y, w) - X(\xi_p) \right)^2. \quad (3.16)$$

Généralement, en vue de l'utilisation ultérieure de la carte, on impose la décorrélation entre  $X(\xi_p)$  et l'erreur  $\widehat{X}_p - X(\xi_p)$ . Plutôt que de chercher à résoudre (3.16), on cherche donc à estimer les poids de *Combinaison Linéaire Interne*.

**Définition 3.6.** On appelle poids de *Combinaison Linéaire Interne* les poids (inconnus)

$$w_p^{ILC} := \arg \min_{w : \mathbf{E} (X(\xi_p)(\widehat{X}_p(Y, w) - X(\xi_p))) = 0} \mathbf{E} \left( \widehat{X}_p(Y, w) - X(\xi_p) \right)^2$$

où  $\widehat{X}_p$  est défini par l'équation (3.15).

Donnons d'autres expressions de  $w_p^{ILC}$ .

$$w_p^{ILC} = \arg \min_{w : \mathbf{E} (X(\xi_p)(\widehat{X}_p(Y, w) - X(\xi_p))) = 0} \mathbf{E} \widehat{X}_p^2(Y, w) - \mathbf{E} X^2(\xi_p) \quad (3.17)$$

$$= \arg \min_{w : w^T \mathbf{1} = 1} \mathbf{E} \widehat{X}_p^2(Y, w) \quad (3.18)$$

$$= \arg \min_{w : w^T \mathbf{1} = 1} w_p^T R_{Y_p} w_p, \quad (3.19)$$

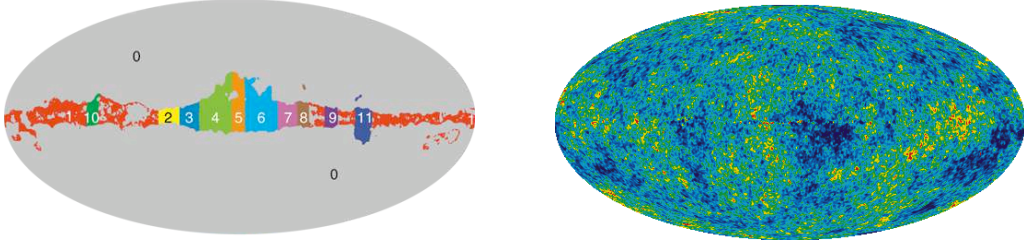


FIG. 3.5 – A gauche : Douze zones sur lesquelles l’estimateur (3.22)–(3.21)–(3.23) est calculé de façon indépendante. A droite : carte estimée  $\widehat{X}^{ILC}$ . Sur chaque zone, cette carte est un barycentre des cinq cartes de la figure 3.4. [images : WMAP]

où  $\mathbf{1}$  est un vecteur de taille  $N_{freq}$  rempli de 1. Le passage de (3.17) à (3.18) se voit en insérant (3.14) dans (3.15).

L’équation (3.19), d’optimisation quadratique en  $w$  sous contrainte linéaire, est résolue grâce aux multiplicateurs de Lagrange, conduisant à

$$w_p^{ILC} = (\mathbf{1}^T R_{Y_p}^{-1} \mathbf{1})^{-1} R_{Y_p}^{-1} \mathbf{1} . \quad (3.20)$$

Si  $\widehat{R}_Y$  est un estimateur de  $R_Y$ , un estimateur possible de  $w_p^{ILC}$  est donc

$$\widehat{w}_p^{ILC} := (\mathbf{1}^T \widehat{R}_{Y_p}^{-1} \mathbf{1})^{-1} \widehat{R}_{Y_p}^{-1} \mathbf{1} . \quad (3.21)$$

**Définition 3.7.** On appelle carte ILC l’ensemble  $\widehat{X}_p$ ,  $1 \leq p \leq N_{pix}$  où

$$\widehat{X}_p := \widehat{X}_p(Y, \widehat{w}_p^{ILC}) \quad (3.22)$$

selon les définitions (3.15) et (3.21).

Le problème à ce stade est donc de trouver un bon estimateur  $\widehat{R}_{Y_p}$ . Si dans (3.14) les avant-plans et le bruit sont, comme le *CMB*, stationnaires, alors  $R_{Y_p} \equiv R_Y$ ,  $1 \leq p \leq N_{pix}$ . Dans ce cas, la covariance empirique

$$\widehat{R}_Y := N_{pix}^{-1} \sum_{p=1}^{N_{pix}} Y_p Y_p^T \quad (3.23)$$

en fournit un estimateur, et les poids  $\widehat{w}^{ILC}$  qui en résultent sont ceux qui minimisent l’équivalent du terme de droite de (3.19) (constant en  $p$ ) où  $R_Y$  a été remplacé par  $\widehat{R}_Y$ . La procédure revient donc à trouver non plus la carte de variance minimale (3.18) mais celle de variance empirique minimale.

En réalité, ni le bruit ni encore moins les avant-plans ne sont stationnaires, comme cela est visible sur la figure 3.4 et a été discuté précédemment. Cependant, ils peuvent être considérés comme *localement stationnaires* sur certaines régions. Les estimateurs précédents sont alors calculés indépendamment pour chaque région, puis “recollés”. La définition de zones lissées et se chevauchant permet d’éviter les effets de bords. C’est exactement la technique employée par l’équipe WMAP pour extraire le *CMB* de leurs données. Leurs 12 zones utilisées, ainsi que le résultat, sont présentés figure 3.5.

## Discussion

Le choix des zones est important et un compromis doit être trouvé quant à leur taille : une taille plus grande améliore l'estimation de  $R_Y$  par  $\widehat{R}_Y$ , mais elle détériore l'approximation de stationnarité locale du bruit et des avant-plans.

Cette approche heuristique a reçu cette année des premiers éléments théoriques d'analyse : l'article de Vio et Andreani [143] ainsi que celui présenté au Chapitre 8 calculent notamment le développement à l'ordre 1 en  $(R_Y - \widehat{R}_Y)$  du biais de l'estimateur (3.22)–(3.21)–(3.23) dans l'asymptotique où  $\widehat{R}_Y \rightarrow R_Y$ , ce qui renseigne quant au choix des zones à faire.

Une remarque importante est que tout ce qui précède est transposable au cas où le modèle (3.14) est exprimé non pas dans le domaine direct, mais dans n'importe quelle autre base. L'implémentation dans le domaine direct présente en effet de sérieux inconvénients, qui justifient l'avertissement des auteurs de la carte 3.5 quant à l'utilisation de cette dernière : elle est visuellement agréable car le *CMB* est estimé sur tout le ciel, mais elle reste contaminée à un niveau important par les avant-plans et par le bruit. En effet, un barycentre des cartes observées est calculé non seulement uniformément en  $p$  sur chaque zone, mais également uniformément à travers les fréquences spatiales  $\ell$ .

Or les propriétés respectives des avant-plans et du bruit varient beaucoup avec la fréquence angulaire  $\ell$ . Intuitivement, les grandes échelles ( $\ell$  petit) sont peu bruitées mais très contaminées, donc très corrélées à travers les longueurs d'onde ( $S_\nu(\xi_p)$  et  $S_{\nu'}(\xi_p)$  proviennent des mêmes avant-plans, vus dans des longueurs d'onde différentes). À l'inverse, les petites échelles ( $\ell$  grand), peu contaminées mais très bruitées, sont très peu corrélées à travers les longueurs d'onde (car le bruit est presque blanc).

La pertinence d'utiliser des ondelettes, dans ce contexte, a été formulée dès 1998 par Tegmark [138], sans que cela ait été repris depuis sous cette forme précise : “la séparation [des composantes] dans le domaine pixels ne tient pas compte de la différence de douceur [spectrale] entre le *CMB* et les avant-plans : la corrélation entre pixels voisins est plus forte pour ces derniers que pour le *CMB* [...] il pourrait donc être préférable de faire la séparation dans le domaine de Fourier. [Mais] les coefficients de séparation doivent aussi dépendre de la position, car le niveau de bruit dépend de la position. Suggérons donc une soustraction des avant-plans par l'utilisation de bases relativement bien localisées tant dans le domaine direct que dans celui de Fourier, par exemple des sortes d'ondelettes.” Une première recherche allant dans ce sens fut proposée en 2003 par Tegmark et al. [139] qui définissent des zones spatiales et des “zones” d'échelles et améliorent la carte 3.5 à partir des mêmes données.

## 3.4 Autres enjeux statistiques

D'autres défis liés à l'étude du Fond diffus cosmologique ne seront pas discutés par la suite mais méritent d'être mentionnés, tant pour leur importance du point de vue physique que pour leur intérêt du point de vue statistique.

### Projection sphérique de données temporelles corrélées

Certains de ces problèmes ne concernent pas l'analyse de données sphériques. Ce sont, aux deux extrémités de la chaîne d'analyse des données *CMB*, le traitement des données temporelles collectées

par les instruments en vue de fabriquer des cartes sphériques ; et l'estimation des paramètres cosmologiques à partir des spectres de puissance angulaire estimés (et conjointement à partir d'autres observations scientifiques indépendantes du *CMB*).

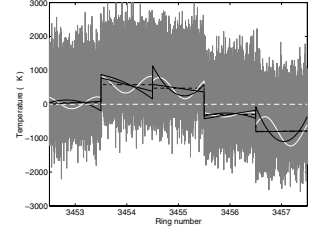
La *fabrication de cartes*, consiste à estimer  $Y_{p,v,d}$  étant données des observations de la forme

$$\tilde{Y}_{t,v,d} = \sum_{p=1}^{N_{pix}} A_{t,p} Y_{p,v,d} + Z_{t,v,d} \quad , \quad (3.24)$$

$$1 \leq t \leq N_{tod}, 1 \leq v \leq N_{freq}, 1 \leq d \leq N_{det}.$$

Pour chaque fréquence lumineuse  $v$  et chaque détecteur  $d$  dans cette fréquence,  $\tilde{Y}_{t,v,d}$ ,  $1 \leq t \leq N_{tod}$ , est la série temporelle des valeurs mesurées le temps de l'expérience ("TOD" est mis pour "time-ordered data"). La matrice  $A$ , dite de pointage, indique l'indice de la direction observée par le détecteur à l'instant  $t$ . Si le bruit  $Z$  est décorrélé entre ces différents indices, l'opération se résume à appliquer au vecteur  $\tilde{Y}_{v,d}$  le pseudo inverse de  $A$  après avoir pondéré par l'inverse du niveau de bruit, autrement dit à attribuer au point  $p$  la valeur moyenne (tenant compte du niveau de bruit) des observations ayant pointé dans cette direction. Notons qu'implicitement cela nécessite d'avoir choisi une pixelisation, en particulier son degré de finesse. Ce dernier résulte d'un compromis à trouver entre la résolution de la carte finale et le niveau de bruit de chaque point.

En réalité, une partie du bruit est corrélée entre les détecteurs et entre les instants de mesures. Il s'agit du bruit à basse fréquence dû à l'électronique de l'instrument et à ses fluctuations thermiques. Une simple projection des données sur la sphère comme expliqué ci-dessus provoque alors des stries le long des lignes de balayages du ciel.



F . 3.6 – Simulation d'un segment de 5h de données PLANCK [image : Keihänen et al. 2004]

## Estimation des paramètres cosmologiques

A l'autre bout de l'analyse et de l'interprétation des données se trouvent la sélection d'un modèle cosmologique et l'estimation des paramètres de ce modèle. Tous les modèles envisagés déterminent une loi jointe de la suite  $C_\ell^X$ . L'estimation, qui consiste à ajuster un spectre paramétrique sur l'ensemble des spectres estimés, est généralement réalisée à travers des simulations.

## Séparation de sources

Pour ce qui est des problèmes posés par des données sphériques, hormis l'extraction du *CMB* et l'estimation spectrale, deux font aujourd'hui l'objet d'un travail intense : l'extraction complète de toutes les composantes du rayonnement mesuré (et non pas seulement le *CMB* comme dans la Section précédent) et les tests de gaussianité et de stationnarité.

En ce qui concerne le problème de séparation de composantes, elle s'exprime sous la forme généralisée suivante du modèle (3.14)

$$Y_{v,p} = X(\xi_p) + \sum_{i=1}^{N_{comp}} S_{v,i}(\xi_p) + \sigma_{v,p} U_{v,p} \quad . \quad (3.25)$$



L'objectif est d'estimer  $X$  et les  $S_{\nu,i}$  (et éventuellement les paramètres du bruit :  $\sigma$ , structure de covariance de  $U$ ). On peut considérer un modèle plus restrictif :

$$S_{\nu,i} = A_{\nu,i} S_i$$

où  $A$  est une matrice, dite de mélange, déterministe, décrivant les spectres lumineux des différentes sources physiques  $S_i$ . Delabrouille et Cardoso [67] font le tour des méthodes de séparation de sources utilisées actuellement, et Leach, Cardoso et al. [100] présentent les performances respectives de certaines méthodes sur des données simulées en vue de préparer l'analyse des données de l'expérience PLANCK.

### Test des hypothèses du modèle (stationnarité, gaussianité)

Enfin, une grande partie de l'analyse statistique du *CMB* repose sur l'hypothèse, justifiable physiquement et validée au moins en première approximation par l'observation, de son caractère stationnaire et gaussien. La qualité des cartes augmentant, il est possible de tester ces hypothèses de manière toujours plus fine, à condition bien sûr de disposer des outils statistiques adéquats. Cabella et Marinucci [174, 155] passent en revue, d'un point de vue statistique, les méthodes actuelles.

## Annexe

Montrons que les équations (3.4) et (3.14) sont justifiées, c'est-à-dire qu'en présence d'avant-plans on observe  $X(\xi) + S_\nu(\xi)$ . Les fluctuations de température sont suffisamment faibles ( $X \sim 10^{-5}$ ) pour pouvoir linéariser la loi de Planck (rappelée page 13) :

$$B(\lambda, T_0(1 + X)) \underset{X \rightarrow 0}{=} B(\lambda, T_0) + \frac{\partial B}{\partial T}(\lambda, T_0)X + o(X).$$

L'additivité des luminances entraîne qu'en présence d'avant-plans, la valeur mesurée de la luminance est

$$L^{obs}(\lambda, T_0) \simeq B(\lambda, T_0) + \frac{\partial B}{\partial T}(\lambda, T_0)X + L^{avant-plans}(\lambda).$$

Les valeurs de  $L(\lambda, T_0)$  et  $\frac{\partial B}{\partial T}(\lambda, T_0)$  étant supposées connues, on peut bien considérer avoir mesuré  $X + S_\nu$  où

$$S_\nu := \left( \frac{\partial B}{\partial T}(T_0, \lambda_\nu) \right)^{-1} L^{avant-plans}(\lambda_\nu)$$

comme annoncé.

## Chapitre 4

# Présentation des résultats obtenus

### Sommaire

4.1	Optimisation non-asymptotique des needlets (chapitre 5) . . . . .	<b>62</b>
	Contexte et motivation . . . . .	62
	Optimisation d'un critère géométrique . . . . .	63
	Optimisation d'un critère statistique . . . . .	63
	Etudes numériques . . . . .	64
4.2	Nouvelle méthode d'estimation spectrale (chapitres 6 et 7) . . . . .	<b>66</b>
	Principe d'estimation . . . . .	66
	Originalité du modèle . . . . .	67
	Convergence des estimateurs . . . . .	68
	Comparaison avec d'autres estimateurs . . . . .	70
	Données longitudinales . . . . .	70
4.3	Applications au Fond diffus cosmologique (chapitres 6 et 8) . . . . .	<b>73</b>
	Estimation spectrale . . . . .	73
	Séparation de sources : extraction du Fond diffus . . . . .	74
4.4	Note technique . . . . .	<b>77</b>
4.5	Perspectives . . . . .	<b>77</b>

Notre objectif est d'utiliser les needlets pour l'analyse statistique du Fond diffus cosmologique, en tirant le meilleur parti de leur localisation simultanée en fréquence (sphérique) et en espace. Pour cela nous optimisons dans un premier temps la localisation spatiale des needlets pour un support fréquentiel donné. Nous utilisons ensuite des needlets pour l'estimation du spectre de puissance angulaire d'un champs sphérique stationnaire, dans un modèle complexe imposé par les conditions pratiques d'observation. Ce modèle comporte un lissage par convolution (qui atténue les hautes fréquences) ; un bruit additif hétéroscédastique ; des données manquantes. L'application de notre estimateur aux données du Fond diffus est complété par une application des needlets à un autre problème statistique lié à ce même Fond diffus : celui de la séparation de source.

## 4.1 Optimisation non-asymptotique des needlets (chapitre 5)

### Contexte et motivation

L'intérêt des needlets d'un point de vue statistique est leur propriété de frame et leur localisation tant spatiale que fréquentielle. Typiquement, cette localisation permet d'analyser un signal aux endroits où il est disponible en étant peu affecté par les données manquantes ou corrompues. Pour certaines applications statistiques (comme la séparation de sources), une propriété de reconstruction exacte est nécessaire. Pour d'autres (comme l'estimation spectrale) elle ne l'est pas.

Quant à la localisation obtenue précédemment (Eq. (2.20)), son utilité est surtout dans l'asymptotique des hautes fréquences,  $j \rightarrow \infty$  (et laisse une liberté certaine dans le choix de la fonction  $\mathbf{a}$ ). Le problème de l'optimisation, non asymptotique, de la localisation des needlets reste, lui, entier. Nous apportons des solutions au choix des filtres assurant les meilleures propriétés de localisation. La figure 4.1 illustre l'impact que peut avoir la forme des filtres fréquentiels sur la localisation spatiale des needlets loin de leur centre.

Pour optimiser la localisation des needlets, nous nous concentrons sur le “design” de ces filtres fréquentiels. Nous commençons par montrer que le schéma dyadique – ou plus généralement  $\mathcal{B}$ -adique au sens des équations (2.13) et (2.15) – peut être assoupli, tout en gardant les propriétés de frame. Cela permet plus de liberté dans la conception des filtres.

Nous considérons ainsi un schéma “biorthogonal”, avec des filtres d'analyse  $(b_{j,\ell})$  band-limited, de fréquence maximale  $\ell_j^{max}$ , non nécessairement  $\mathcal{B}$ -adiques, et des filtres de synthèse

$$\tilde{b}_{j,\ell} := \left( \sum_{j'} b_{j',\ell}^2 \right)^{-1} b_{j,\ell} .$$

Les needlets d'analyse et de synthèse sont respectivement définies par

$$\psi_{j,k}(\xi) := \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\infty} b_{j,\ell} L_{\ell}(\xi \cdot \xi_{j,k}) \quad (4.1)$$

et

$$\tilde{\psi}_{j,k}(\xi) := \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\infty} \tilde{b}_{j,\ell} L_{\ell}(\xi \cdot \xi_{j,k}) ,$$

où  $(\xi_{j,k}, \lambda_{j,k})_{k \in \{1, \dots, N_j\}}$  assure une cubature à l'ordre  $2\ell_j^{max}$ . Quant aux coefficients de needlet d'une fonction  $f$ , ce sont toujours  $\beta_{j,k} := \langle f, \psi_{j,k} \rangle$  et  $\tilde{\beta}_{j,k} := \langle f, \tilde{\psi}_{j,k} \rangle$ .

**Proposition.** *S'il existe des constantes strictement positives  $A$  et  $B$  telles que*

$$\forall \ell \in \mathbb{N}, \quad A \leq \sum_{j \in \mathbb{N}} |b_{j,\ell}|^2 \leq B ,$$

*alors  $(\psi_{j,k})_{j \in \mathbb{N}, k \in \{1, \dots, N_j\}}$  est une frame de  $\mathbb{L}^2(\mathbb{S}^2)$  dont  $A$  et  $B$  sont les bornes, et  $\tilde{\psi}_{j,k}$  est sa frame duale.*

## Optimisation d'un critère géométrique

L'utilité de la localisation spatiale des needlets est pour nous l'insensibilité de leurs coefficients à un masque appliqué à  $f$ , suffisamment loin de ce masque (ainsi, comme on le verra, que la possibilité qu'elles offrent de s'adapter au niveau local d'un bruit). De ce point de vue, le critère de localisation le plus approprié est la concentration dans une région donnée, que l'on choisira être une calotte sphérique (boule pour la distance  $d$ ) :

$$C_{p,r}(\psi) := \frac{\left\{ \int_{\mathbb{S}^2} |\psi(\xi)|^p \mathbf{I}(d(\xi, N) \leq r) d\xi \right\}^{1/p}}{\left\{ \int_{\mathbb{S}^2} |\psi(\xi)|^p d\xi \right\}^{1/p}}.$$

D'un point de vue physique, on s'intéresse généralement plutôt à la quantité  $1 - C_{p,r}(\psi)$  qui mesure la capacité de réjection de  $\psi$  à l'égard de "pollutions" hors de sa zone centrale (la calotte sphérique de rayon  $r$ ).

**Proposition.** Soient une suite  $(r_j)_{j \in \mathbb{N}}$  dans  $[0, \pi[$  et des suites  $(\ell_j^{\min})_{j \in \mathbb{N}}$  et  $(\ell_j^{\max})_{j \in \mathbb{N}}$  dans  $\mathbb{N}$  telles que  $\ell_j^{\min} < \ell_j^{\max}$  ( $\ell_j^{\min}$  et  $\ell_j^{\max}$  sont les fréquences minimale et maximale des needlets de la bande  $j$ ). Alors, si  $\psi_{j,k}$  est définie par l'équation (4.1),

$$\forall j \in \mathbb{N}, \forall k \in \{1, \dots, N_j\}, C_{2,r_j}(\psi_{j,k}) = \frac{\bar{b}_j^T \mathbf{D} \bar{b}_j}{\|\bar{b}_j\|^2}$$

où  $\bar{b}_j$  le vecteur  $(\bar{b}_j)_\ell := \sqrt{\frac{2\ell+1}{8\pi^2}} b_{j,\ell}$ ,  $\ell_j^{\min} \leq \ell \leq \ell_j^{\max}$ , et  $\mathbf{D}_{\ell,\ell'} := \frac{8\pi^2}{\sqrt{(2\ell+1)(2\ell'+1)}} \int_{\cos r_j}^1 L_\ell(z) L_{\ell'}(z) dz$ .

Cela nous permet, pour chaque bande  $j$ , de maximiser le critère  $C_{2,r_j}$  parmi les needlet dont le support fréquentiel est fini, inclus dans  $[\ell_j^{\min}, \ell_j^{\max}]$ . Cela revient au calcul du vecteur propre de la matrice  $\mathbf{D}$  associé à la plus grande valeur propre.

## Optimisation d'un critère statistique

Le second type de critère envisagé repose sur l'idée suivante : puisque dans certaines applications, on n'utilise la localisation spatiale qu'à travers ses conséquences en termes statistiques, pourquoi ne pas optimiser directement un critère "statistique" ? Nous nous posons le problème suivant : étant donné un champs aléatoire  $X$  à analyser à partir de l'observation de  $WX$  où  $W \in \mathbb{L}^2(\mathbb{S}^2)$  est une fonction déterministe connue (par exemple l'indicatrice d'une région observée), quelle needlet a ses coefficients les moins affectés par la présence de  $W$ , au moins dans une certaine région d'intérêt (par exemple là où  $W = 1$ ) ?

Plus précisément, soient

$$\epsilon_{j,k} := \frac{\beta_{j,k} - \beta'_{j,k}}{\sqrt{\mathbf{E}(\beta'_{j,k})^2}}$$

les erreurs normalisées dues à  $W$ , où  $\beta_{j,k} := \langle WX, \psi_{j,k} \rangle$  sont les coefficients observés de  $WX$  et  $\beta'_{j,k} := \langle X, \psi_{j,k} \rangle$  sont les coefficients inconnus de  $X$ .

**Lemme.** On a  $\epsilon_{j,k} = \epsilon(\xi_{j,k})$  où, avec  $\Psi_j := \sum_{\ell=0}^{\infty} b_{j,\ell} L_\ell$ ,

$$\epsilon(\xi) := \left( \mathbf{E} \left( (\Psi_j * X)(\xi) \right)^2 \right)^{-1/2} (\Psi_j * X - \Psi_j * (WX))(\xi).$$

Nous minimisons alors l'erreur intégrée dans une région, étant donnée la fonctions  $W$ , en montrons que cette erreur intégrée est elle aussi une forme quadratique des filtres. Soit en effet une fonction  $D \in \mathbb{L}^2(\mathbb{S}^2)$  (par exemple l'indicatrice d'une région d'intérêt), et soit

$$R(b, W, D, X) := \mathbf{E} \int_{\mathbb{S}^2} D(\xi) |\epsilon(\xi)|^2 d\xi$$

où  $X$  est mis pour "la loi de  $X$ " et  $b := (b_\ell)_{\ell \in \mathbb{N}}$ .

**Proposition.** Si  $X$  est gaussien, centré, stationnaire de spectre de puissance  $(C_\ell)_{\ell \in \mathbb{N}}$ , alors

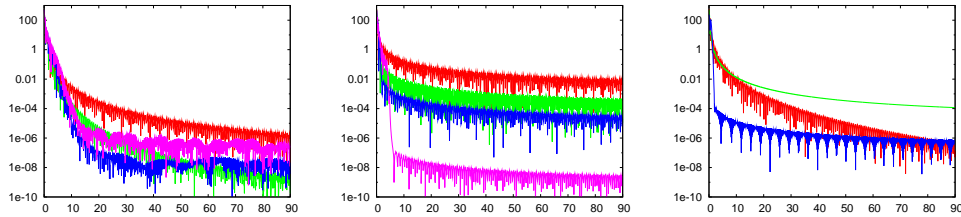
$$R(b, W, D, X) = \frac{b^T \mathbf{Q} b}{\|b\|^2}$$

où

$$\mathbf{Q}_{\ell, \ell'} := \sum_{\ell_1} C_{\ell_1} \sum_{\ell_1, \ell_2, \ell_3, \ell_4, m_1, m_2, m_3, m_4, m, m'} a_{\ell_2, m_2}^{1-W} (a_{\ell_3, m_3}^{1-W})^* a_{\ell_4, m_4}^D \times \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m_1 & m_3 & m' \end{bmatrix} \begin{bmatrix} \ell & \ell_4 & \ell' \\ m & m_4 & m' \end{bmatrix}$$

$$\text{avec } \begin{bmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{bmatrix} := \int_{\mathbb{S}^2} \mathcal{Y}_{\ell, m}(\xi) \mathcal{Y}_{\ell', m'}(\xi) \mathcal{Y}_{\ell'', m''}^*(\xi) d\xi.$$

Ainsi, comme dans le cas du critère géométrique pour la norme  $\mathbb{L}^2$ , la minimisation du critère statistique  $R(b)$  se réduit au calcul du vecteur propre associé à la plus petite valeur propre de  $\mathbf{Q}$ .



F . 4.1 – Fonctions  $\theta \mapsto \Psi(\cos \theta) := \sum_{\ell=0}^{\infty} b_\ell L_\ell(\cos \theta)$  pour différents choix de filtres  $b_\ell$ . Tous les filtres (sauf dans l'image de droite) sont limités à la même bande de fréquence. L'échelle logarithmique fait ressortir des caractéristiques significativement différentes de la décroissance de  $\Psi$ . Ces différences peuvent avoir un impact non négligeable en pratique, par exemple en présence de données manquantes ou corrompues.

## Etudes numériques

Nous avons mis en oeuvre numériquement les idées précédentes et comparé les différentes needlets obtenues. Tout d'abord, la réalisation pratique de la transformation en needlet (c'est-à-dire le calcul des coefficients  $\beta_{j,k}$ ) est effectuée via des transformations en Harmoniques Sphériques directe et

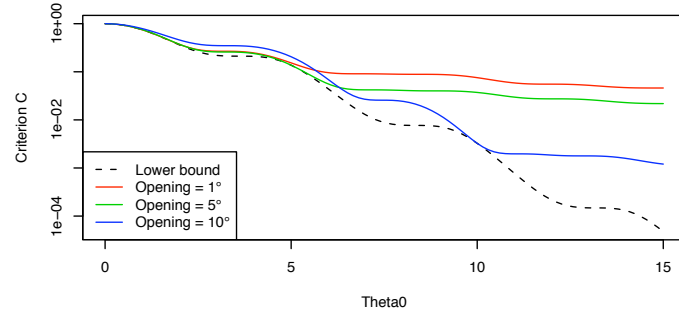


FIG. 4.2 – Fonctions  $r \mapsto 1 - C_{2,r}(\psi)$  pour trois fonctions  $\psi$  à support fréquentiel [33, 64], réalisant le maximum du critère  $C_{2,r}$  pour trois valeurs différentes de  $r$  (ici  $r$  est noté “Theta0”). Plus le rayon  $r$  dans lequel on impose la concentration est petit, moins le niveau de réjection ( $1 - C_{2,r}$ ) est bas loin du centre.

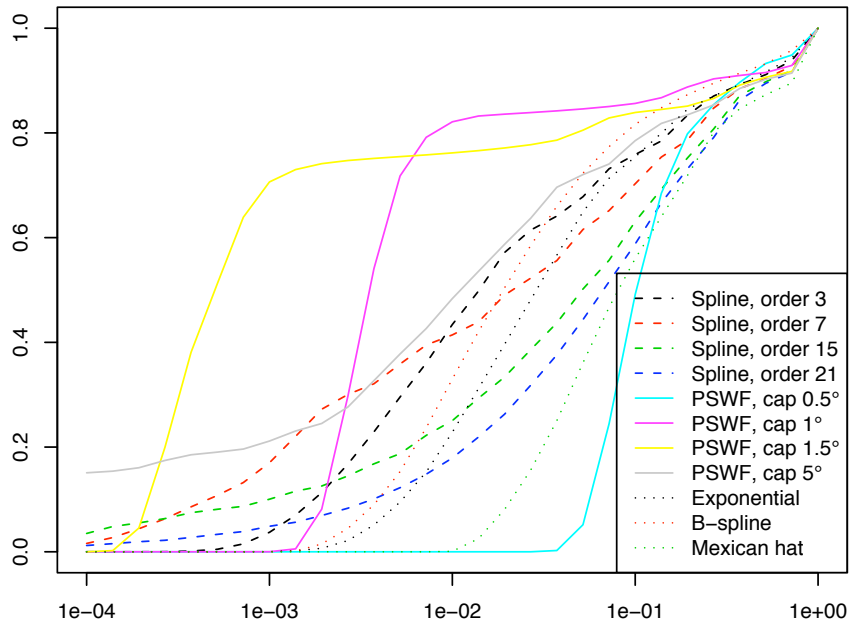


FIG. 4.3 – Pour un masque  $W$  à valeurs dans  $\{0, 1\}$ , et différents filtres (voir le Chapitre 5), fonctions de répartition empirique des erreurs  $\epsilon_{j,k}$  dans la région  $\{W = 1\}$ .

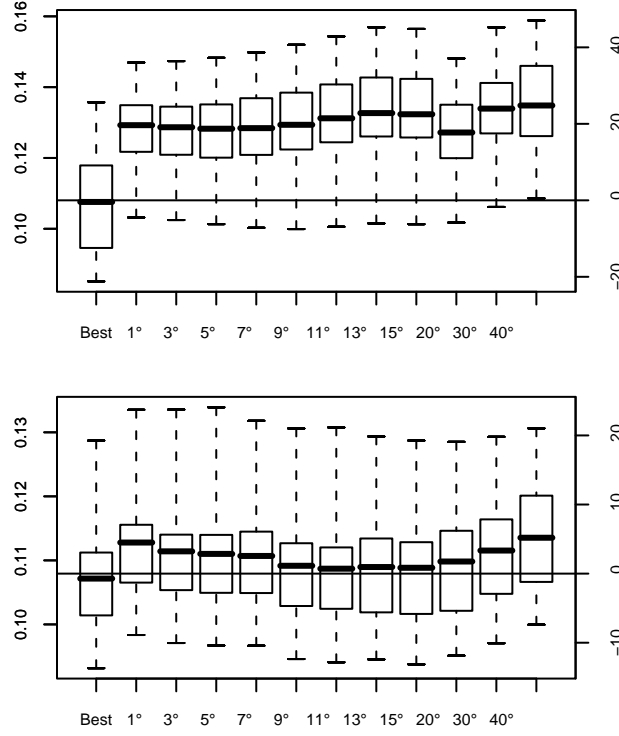


FIG. 4.4 – Distribution empirique de l’erreur  $(-\epsilon_{j,k})$  pour une bande  $j$  fixée et différentes needlelets (celle – notée “Best” – optimisant le critère statistique  $R(b)$  et celles optimisant le critère géométrique  $C_{\Omega_j}$  pour des boules de différents rayons). Les deux graphiques correspondent à deux masques  $W$  différents.

inverse, et un filtrage dans le domaine de Fourier, selon le schéma

$$(f(\xi_P))_{1 \leq p \leq N} \xrightarrow{\text{SHT}} (a_{\ell,m})_{\ell \leq \ell_j^{\max}} \xrightarrow{\times} (b_{j,\ell} a_{\ell,m})_{j \leq J, \ell \leq \ell_j^{\max}} \xrightarrow{\text{SHT}^{-1}} \left( \lambda_{j,k}^{-\frac{1}{2}} \beta_{j,k} \right)_{j \leq J, 1 \leq k \leq N_j}.$$

Cela résulte de l’observation que les coefficients de needlelet  $\beta_{j,k}$  sont les échantillonnés de la fonction à analyser lissée par les filtres. Alors que les coefficients n’auraient pas pu être calculés en autant d’intégrales numériques du produit  $f\psi_{j,k}$ , cette méthode permet leur calcul rapide.

Nous avons comparé numériquement les needlelets optimisées pour les différents critères présentés, ainsi que d’autres couramment utilisées (comme le “chapeau mexicain”). Les graphiques 4.3 et 4.4, ainsi que le tableau 5.8 page 96, illustrent certains de ces résultats. Il ressort en particulier que le gain apporté par l’optimisation du critère “statistique” est dans certaines situations important, au regard de ce critère, même par rapport aux needlelets optimisant un critère de concentration.

## 4.2 Nouvelle méthode d’estimation spectrale (chapitres 6 et 7)

### Principe d’estimation

La principale utilisation statistique que nous avons faite des propriétés des needlelets est l’introduction d’une nouvelle méthode d’estimation du spectre de puissance d’un champs sphérique stationnaire

centré  $X$ . L'importance de ce spectre, qui détermine la loi du champs dans le cas gaussien, a été soulignée dans le premier chapitre.

L'estimation à partir des coefficients de needlets est fondée sur l'observation que la variance de ceux-ci, dans le cas idéal de l'observation parfaite de  $X$ , est une version lissée du spectre. Plus précisément, la covariance entre les coefficients de la bande  $j$  aux points  $\xi_{j,k}$  et  $\xi_{j,k'}$  vaut

$$\mathbf{E} \beta_{j,k} \beta_{j,k'} = \sqrt{\lambda_{j,k} \lambda_{j,k'}} \sum_{\ell=0}^{\infty} b_{j,\ell}^2 C_{\ell} L_{\ell}(\xi_{j,k} \cdot \xi_{j,k'}).$$

De plus, en présence d'un bruit hétéroscédastique de niveau connu, nous avons déterminé la structure de covariance de l'influence de ce bruit sur les coefficients de needlet.

**Proposition.** *Si l'observation des valeurs  $X(\xi_{j,k})$  d'un champ est entachée d'une erreur  $\sigma_{j,k} U_{j,k}$  où les variables  $U_{j,k}$  sont décorréliées, centrées et réduites, alors les erreurs  $\zeta_{j,k}$  sur les coefficients  $\beta_{j,k}$  sont centrées et de covariance*

$$\mathbf{E} \zeta_{j,k} \zeta_{j,k'} = \frac{1}{\sqrt{\lambda_{j,k} \lambda_{j,k'}}} \sum_{p=1}^{N_j} (\lambda_{j,p} \sigma_{j,p})^2 \psi_{j,k}(\xi_{j,p}) \psi_{j,k'}(\xi_{j,p}).$$

Dans chaque bande  $j$  on peut donc estimer le spectre lissé par une moyenne des carrés des coefficients de needlet, débiaisés du bruit. L'intérêt de cette approche est que grâce à la localisation des coefficients il est possible de contrôler l'influence de données manquantes sur l'estimateur et d'adapter les poids de la moyenne au niveau local du rapport signal/bruit.

## Originalité du modèle

Précisons quelques notations. On considère le modèle suivant (plus de détails sont donnés dans le chapitre 7) : on observe une suite d'expériences, indexées par  $j$  :

$$Y_{j,p} = W_{j,p} \left[ (B_j * X)(\xi_{j,p}) + \sigma_{j,p} U_{j,p} \right] \quad , \quad p \in \{1, \dots, N_j\}. \quad (4.2)$$

Les fonctions  $W_{j,\cdot}$ ,  $B_j$  et  $\sigma_{j,\cdot}$  sont déterministes et connues. Les  $W_{j,p}$  sont supposés positifs et uniformément bornés (typiquement,  $W_{j,p} = 1$  si le point  $\xi_{j,p}$  est observé, 0 sinon). La fonction  $B_j \in \mathbb{L}^2([-1, 1])$  vérifie  $(B_j)_{\ell}^{\wedge} = 1$  si  $\ell \leq \ell_j^{\max}$  et 0 si  $\ell > 2 \ell_j^{\max}$  (la convolution par  $B_j$  a donc un effet “passe-bas”). Les variables  $U_{j,p}$ ,  $j \in \mathbb{N}$ ,  $p \in \{1, \dots, N_j\}$ , sont  $\stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ .

Les paramètres à estimer sont les  $C^{(j)}$ ,  $j \in \mathbb{N}$ , définis par

$$C^{(j)} := (4\pi)^{-1} \sum_{\ell=0}^{\infty} (2\ell + 1) b_{j,\ell}^2 C_{\ell}.$$

L'originalité du modèle considéré consiste en

- L'observation d'une réalisation *unique* du champs  $X$ .
- La *complexité* du modèle, incluant une convolution du signal d'intérêt, des données manquantes et un bruit hétéroscédastique additif.
- La prise en compte d'observations *discrètes*, échantillonnées sur la sphère.

La raison de ces choix est de disposer d'un modèle suffisamment réaliste pour pouvoir appliquer notre méthode aux données des expériences d'observations du CMB.



Du fait de l'unicité de la réalisation de  $X$ , il ne peut pas être question de l'asymptotique habituelle d'un nombre d'observations (souvent i.i.d.) tendant vers l'infini. Par contre, comme cela a été discuté dans le chapitre 3 (équation (3.1)), il est possible de considérer l'asymptotique des hautes fréquences. Cela nécessite une définition particulière de la convergence d'estimateurs.

**Définition 4.1.** *Étant donnée une suite de paramètres non nuls  $\theta_j$ ,  $j \in \mathbb{N}$ , et une suite d'estimateurs  $\widehat{\theta}_j$  de  $\theta_j$ , nous dirons que  $(\widehat{\theta}_j)$  est convergente si*

$$\frac{\mathbf{E} (\widehat{\theta}_j - \theta_j)^2}{\theta_j^2} \xrightarrow{j \rightarrow \infty} 0. \quad (4.3)$$

### Convergence des estimateurs

La bonne localisation des needlets nous permet de sélectionner des coefficients peu affectés par les données manquantes. Pour cela, donnons-nous une suite de seuils  $t_j > 0$ ,  $j \in \mathbb{N}$ , et définissons les ensembles déterministes suivants, correspondant aux coefficients conservés.

$$\mathcal{K}_j := \left\{ k \in \{1, \dots, N_j\} : \left\{ \sum_{p=1}^{N_j} \lambda_{j,p} (1 - W_{j,p})^2 (\psi_{j,k}(\xi_{j,p}))^2 \right\}^{1/2} \leq t_j \right\}.$$

Les estimateurs proposés sont de la forme

$$\widehat{C}^{(j)} := \sum_{k=1}^{N_j} \mathbf{I}(k \in \mathcal{K}_j) w_{j,k} \left[ \lambda_{j,k}^{-1} \beta_{j,k}^2 - n_{j,k}^2 \right] \quad (4.4)$$

où les  $\beta_{j,k}$ ,  $k \in \{1, \dots, N_j\}$  sont les coefficients de needlet du champ discret observé  $Y_{j,\cdot}$ ,  $n_{j,k}^2 := \lambda_{j,k}^{-1} \mathbf{E} \zeta_{j,k}^2$  sont les termes de débiaisage du bruit et  $w_{j,k}$  sont des poids de mélange vérifiant

$$\sum_{k=1}^{N_j} \mathbf{I}(k \in \mathcal{K}_j) w_{j,k} = 1.$$

Nous obtenons deux résultats de convergence, dans le cas où les needlets sont construites selon le schéma  $\mathcal{B}$ -adique des équations (2.15) et (2.16) avec une fonction  $\mathbf{a}$  de classe  $C^M$ ,  $M$  assez grand. Le premier résultat ne suppose (presque) aucune connaissance a priori sur le spectre et considère des poids  $w_{j,k}$  très généraux. Le second suppose connu l'ordre de grandeur du spectre, donc du rapport signal/bruit, et utilise des poids  $w_{j,k}$  qui tiennent compte de ce rapport *local*. Cette hypothèse, conforme au demeurant à la situation pratique qui se pose pour l'analyse du CMB, permet d'abaisser fortement les conditions suffisantes concernant le bruit pour obtenir la convergence.

Dans les deux cas, une hypothèse de régularité est faite sur le spectre à estimer : on suppose, que pour  $\ell \in [\mathcal{B}^{j-1}, \mathcal{B}^{j+1}]$ ,  $C_\ell = \ell^{-\alpha} \mathbf{g}_j(\mathcal{B}^{-j}\ell)$  où les  $\mathbf{g}_j$  ont uniformément leur dérivées successives bornées. On suppose également connaître une borne inférieure de  $\alpha$ , de façon à choisir des seuils  $t_j$  vérifiant  $t_j = o(\mathcal{B}^{-\alpha j})$ .

Pour énoncer le premier résultat, convenons d'appeler *asymptotiquement mélangeants* des poids  $w_{j,k}$  vérifiant

$$\sum_{k,k'=1}^{N_j} \mathbf{I}\left((k,k') \in \mathcal{K}_j^2\right) (1 + \mathcal{B}^j d(\xi_{j,k}, \xi_{j,k'}))^{-M} |w_{j,k} w_{j,k'}| \xrightarrow{j \rightarrow \infty} 0$$

et

$$\sup_{1 \leq p \leq N_j} \sum_{k,k'=1}^{N_j} \frac{\mathbf{I}\left((k,k') \in \mathcal{K}_j^2\right) |w_{j,k} w_{j,k'}|}{(1 + \mathcal{B}^j d(\xi_{j,p}, \xi_{j,k}))^{2M} (1 + \mathcal{B}^j d(\xi_{j,p}, \xi_{j,k'}))^{2M}} \xrightarrow{j \rightarrow \infty} O(\mathcal{B}^{-4j}).$$

Par exemple, des poids vérifiant  $\max_{k \in \mathcal{K}_j} |w_{j,k}| \xrightarrow{j \rightarrow \infty} O(N_j^{-1})$  sont asymptotiquement mélangeants.

**Théorème.** *Si les poids  $w_{j,k}$  sont asymptotiquement mélangeants et si le niveau de bruit vérifie*

$$\sum_{p=1}^{N_j} \lambda_{j,p} (W_{j,p} \sigma_{j,p})^4 \xrightarrow{j \rightarrow \infty} o(\mathcal{B}^{(4-\alpha)j}), \quad (4.5)$$

*alors la suite d'estimateurs  $\widehat{C}_j$  de  $C^{(j)}$  est convergente.*

Si maintenant un ordre de grandeur  $\overline{C}^{(j)}$  de  $C^{(j)}$  est connu, alors les poids  $w_{j,k}$  sont définis en s'inspirant de l'observation suivante.

**Lemme.** *Dans le modèle  $Y_k = X_k + n_k U_k$ ,  $k \in \{1, \dots, N\}$ ,  $X_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, C)$ ,  $U_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ , le paramètre  $C^*$  maximisant la vraisemblance de  $C$  vérifie  $C^* = \sum_{k=1}^N w_k(C^*, n) (Y_k^2 - n_k^2)$  où  $w_k(C, n) := (C + n_k^2)^{-2} \left[ \sum_{k'=1}^N (C + n_{k'}^2)^{-2} \right]^{-1}$ .*

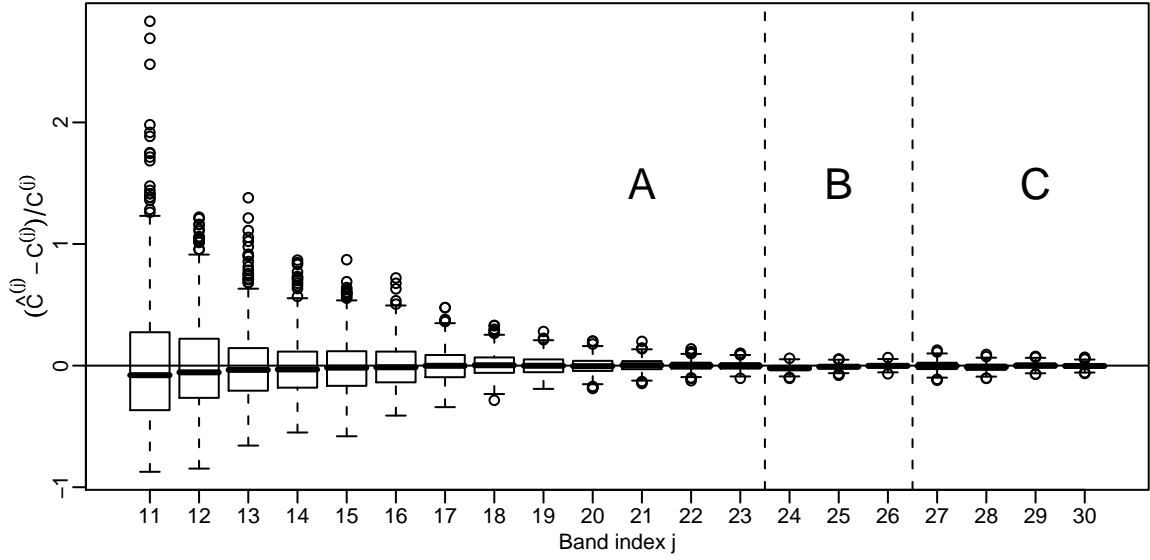
Bien que les coefficients  $\beta_{j,k}$  ne soient pas i.i.d., leur décorrélation asymptotique suggère donc d'utiliser les estimateurs (4.4) avec les poids

$$w_{j,k} = (C + n_{j,k}^2)^{-2} \left[ \sum_{k' \in \mathcal{K}_j} \left( \overline{C}^{(j)} + n_{j,k'}^2 \right)^{-2} \right]^{-1}. \quad (4.6)$$

Grâce à ces poids, la condition (4.5) peut être remplacée par une condition autorisant un niveau de bruit arbitrairement élevé (et même infini) dans certaines régions, pourvu que dans certaines autres le rapport signal/bruit soit contrôlé. Cela peut être quantifié par la condition

$$\exists \kappa > 1, \exists (r_j) \in (\mathbb{R}_+)^{\mathbb{N}}, \frac{(1 + r_j)^2}{\text{Card} \{k \in \mathcal{K}_j : (n_{j,k})^2 \leq r_j C^{(j)}\}} \xrightarrow{j \rightarrow \infty} O(\mathcal{B}^{-\kappa j}). \quad (4.7)$$

**Théorème.** *Si les poids sont choisis selon (4.6), s'il existe une constante  $c > 0$  telle que  $\forall j \in \mathbb{N}$ ,  $c^{-1} C^{(j)} \leq \overline{C}^{(j)} \leq c C^{(j)}$  et si la condition (4.7) est vérifiée, alors la suite d'estimateurs  $\widehat{C}_j$  de  $C^{(j)}$  est convergente.*



F . 4.5 – Distribution empirique des quantités  $\frac{\widehat{C}_j - C_j}{C_j}$ .

### Comparaison avec d'autres estimateurs

Nous avons illustré le comportement de ces derniers estimateurs par plusieurs simulations. Dans une première, décrite Section 7.4, le schéma  $\mathcal{B}$ -adique ayant servi à l'obtention des résultats asymptotiques est respecté. La distribution empirique du terme de droite dans l'équation (4.3), illustrée par la figure 4.5, est conforme aux attentes.

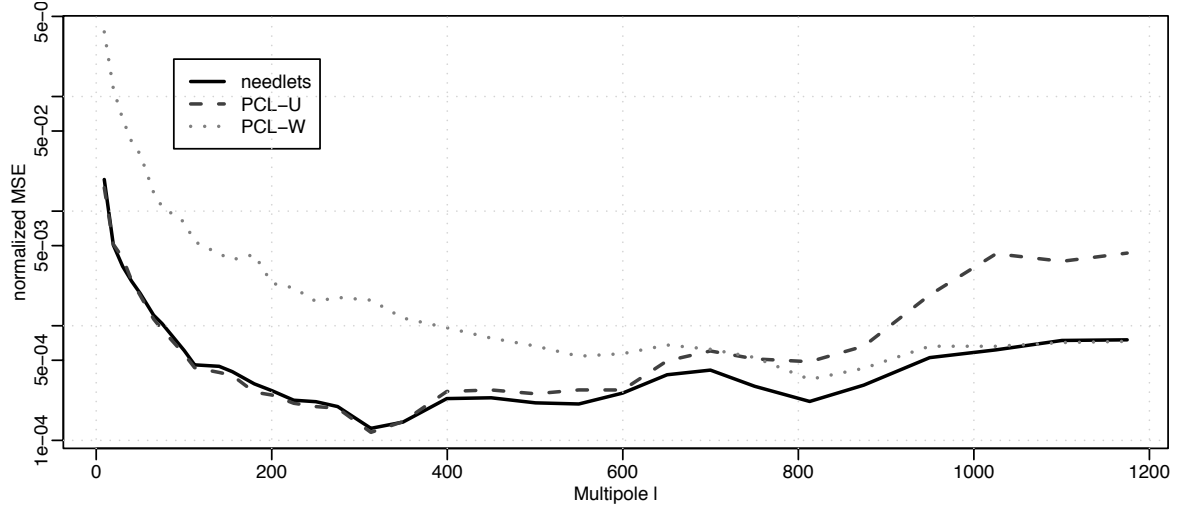
Dans une seconde expérience numérique, décrite Section 6.3.1, les bandes de fréquences  $[\ell_j^{\min}, \ell_j^{\max}]$  considérées sont quelconques. Nous comparons les performances de nos estimateurs à celles des estimateurs  $\widehat{C}_\ell^{PCL}(Y, 1)$  et  $\widehat{C}_\ell^{PCL}(Y, \sigma^{-2})$  introduits au chapitre 3 (équation (3.9)). Ce sont les estimateurs utilisés le plus couramment pour l'estimation spectrale du CMB.  $\widehat{C}_\ell^{PCL}(Y, 1)$  est utilisé pour les basses fréquences et  $\widehat{C}_\ell^{PCL}(Y, \sigma^{-2})$  pour les hautes fréquences. Le passage de l'un à l'autre est fait à une fréquence déduite de simulations, et parfois une transition permet de passer de l'un à l'autre (voir le chapitre 6 pour plus de détails et des références).

La comparaison entre les trois estimateurs (figure 4.6) montre que notre estimateur est presque uniformément (pour toutes les fréquences) meilleur que le meilleur des deux autres. Surtout, il est adaptatif au sens où il assure automatiquement la transition entre le régime des basses fréquences peu bruitées et celui des hautes fréquences très bruitées.

### Données longitudinales

Un point fort de l'estimation spectrale à partir des coefficients de needlets, outre l'adaptabilité au niveau local du bruit, est qu'elle permet l'agrégation de plusieurs observations simultanées dans chaque bande de fréquence. Considérons la généralisation suivante du modèle (4.2) :

$$Y_{j,p,e} = W_{j,p,e} \left[ (B_{j,e} * X)(\xi_{j,p}) + \sigma_{j,p,e} U_{j,p,e} \right] \quad , \quad p \in \{1, \dots, N_j\} \quad , \quad e \in \{1, \dots, E_j\} .$$



F . 4.6 – Comparaison de l’estimateur  $\widehat{C}^{(j)}$  défini par les équations (4.4)–(4.6) et noté “needlets” avec les estimateurs  $\widehat{C}^{(j),PCL}(Y, 1)$  (noté “PCL-U”) et  $\widehat{C}^{(j),PCL}(Y, \sigma^{-2})$  (noté “PCL-W”) définis par l’équation (3.9). Comme attendu, les performances des estimateurs “PCL-U” et “PCL-W” se croisent, l’un étant meilleur aux basses fréquences peu bruitées, l’autre aux hautes fréquences très bruitées. L’estimateur “needlets” est presque uniformément meilleur sur l’ensemble des fréquences.

Ce modèle est celui de données longitudinales, dans la mesure où  $N_j \gg E_j$  ( $E_j$  correspond physiquement au nombre d’expériences ou de détecteurs indépendants réalisant des observations). Le même signal d’intérêt  $X$  est observé avec des bruits indépendants et avec différentes fonctions  $W$ ,  $B$  et  $\sigma$ .

Dans la section 6.3.2 nous réalisons des simulations comportant des observations d’une large fraction de la sphère avec un bruit élevé, et des observations de petites régions avec un niveau de bruit plus faible. Cela est illustré par la figure 4.7(a), et correspond à la réalité physique des observations.

Dans le cas où un ordre de grandeur  $\overline{C}^{(j)}$  de  $C^{(j)}$  on utilise la généralisation suivante du dernier lemme.

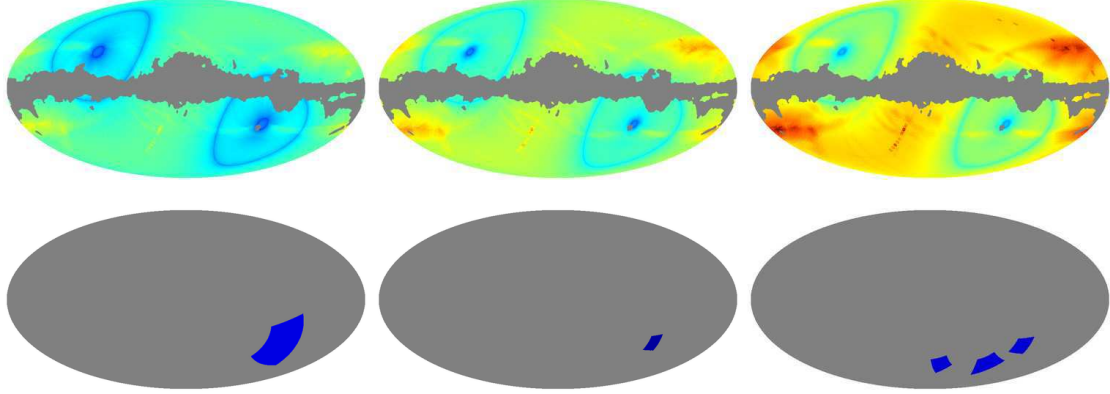
**Lemme.** Dans le modèle  $Y_{k,e} = X_k + n_{k,e}U_{k,e}$ ,  $k \in \{1, \dots, N\}$ ,  $e \in \{1, \dots, E\}$ ,  $X_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, C)$ ,  $U_{k,e} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$ , le paramètre  $C^*$  maximisant la vraisemblance de  $C$  vérifie

$$C^* = \sum_{k=1}^N w_k(C^*, n) \left( \left( \sum_{e=1}^E \omega_{k,e}(n) Y_{k,e} \right)^2 - \widetilde{n}_k^2 \right) \quad (4.8)$$

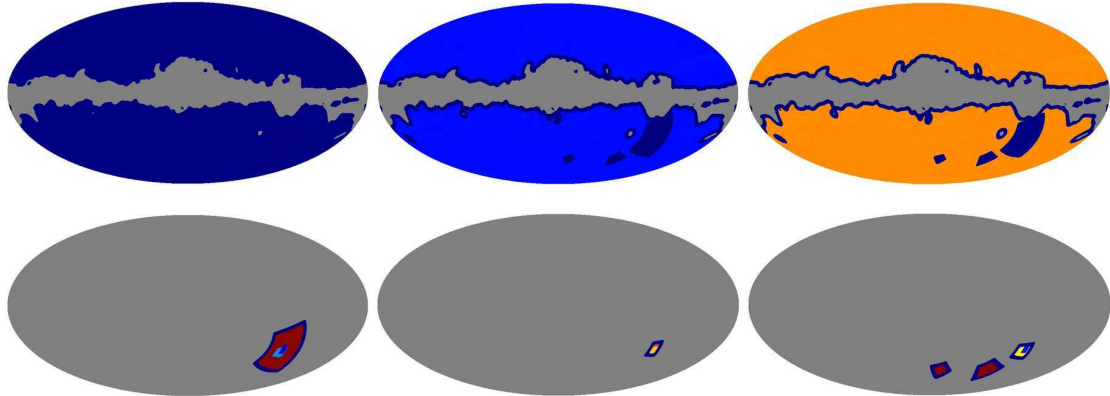
où  $\omega_{k,e}(n) := n_{k,e}^{-2} \left[ \sum_{e=1}^E n_{k,e}^{-2} \right]^{-1}$ ,  $\widetilde{n}_k = \left( \sum_{e=1}^E n_{k,e}^{-2} \right)^{-1/2}$  et  $w_k(C, n) := \left( C + \widetilde{n}_k^2 \right)^{-2} \left[ \sum_{k'=1}^N \left( C + \widetilde{n}_{k'}^2 \right)^{-2} \right]^{-1}$ .

Le terme de droite de l’équation (4.8) fait d’abord intervenir une moyenne en  $e$ , pour les poids  $\omega_{k,e}$  des observations  $Y_{k,e}$ . Cela correspond à une fusion des expériences physiques. Cela est suivi (après élévation au carré et débiaisage) d’une moyenne en  $k$  selon les poids  $w_k$  qui dépendent du rapport signal/bruit local pour un “niveau de bruit équivalent”  $\widetilde{n}$  dans la carte agrégée.

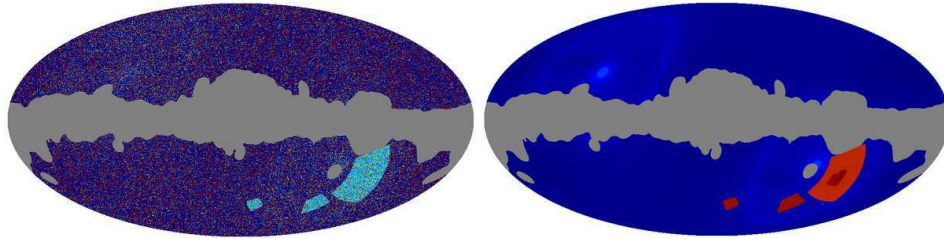
Notre estimateur de  $C^{(j)}$  est obtenu en insérant, dans le terme de droite de l’équation (4.8), les coefficients  $\lambda_{j,k}^{-1/2} \beta_{j,k,e}$  au lieu de  $Y_{k,e}$  et  $\overline{C}^{(j)}$  au lieu de  $C^*$ . Des exemples de cartes de poids  $\omega_{j,\cdot,e}$  et  $W_{j,\cdot}$  ainsi obtenues sont représentées figure 4.7(b-c). Les résultats de cette procédure sur des simulations sont



(a) Niveaux de bruit “masqués” (i.e. on représente  $\sigma W$ ).

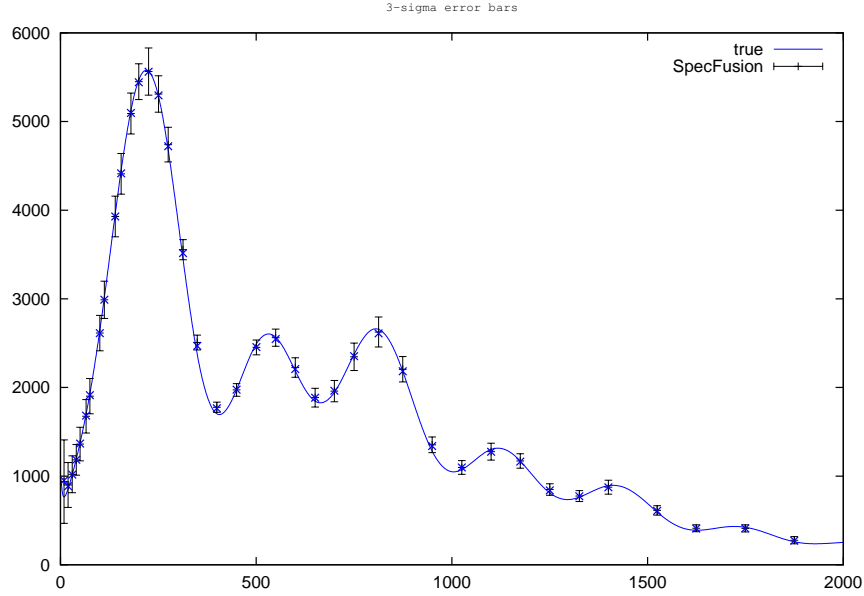


(b) Poids  $\omega_{j,k,e}$  affectés localement aux différentes expériences, dans une bande  $j$  donnée.



(c) A gauche : pour la même bande  $j$ , carte des coefficients de needlet “agrégés”  $\sum_{e=1}^E \omega_{j,k,e} \lambda_{j,k}^{-1} \beta_{j,k,e}$ . A droite : poids  $w_{j,k}$  moyennant la carte de gauche pour estimer  $C^{(j)}$ .

F . 4.7 – Illustration de la méthode d’estimation dans le cas de données longitudinales.



F . 4.8 – Résultats de simulations pour l’estimation de  $C^{(j)}$ . Sont représentées : la courbe des valeurs de  $C_\ell$  définissant les  $C^{(j)}$  ; les valeurs des paramètres  $C^{(j)}$  à estimer (croix bleues) et les valeurs moyennes et barres d’erreur de nos estimateurs (en noir). La demi-longueur des barres d’erreur est égale à 3 écarts-types de la distribution empirique des estimateurs. Cette estimation du spectre à partir de plusieurs expériences est à comparer à la fig. 3.3 page 55.

donnés par la figure 4.8. Ces résultats numériques, pour ces estimateurs qui sont les seuls à pouvoir agréger l’ensemble des données, sont prometteurs.

### 4.3 Applications au Fond diffus cosmologique (chapitres 6 et 8)

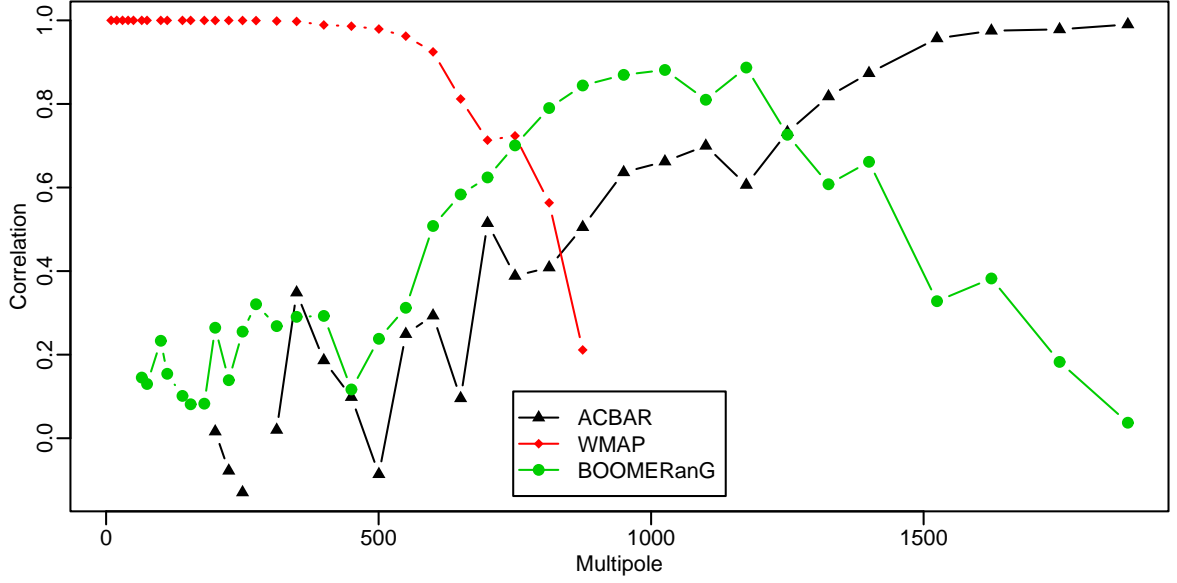
#### Estimation spectrale

La méthode d’estimation présentée dans la section précédente a été appliquée à des simulations réalistes d’observations du CMB. Dans le cas de la simulation d’une seule expérience, il apparaît que notre méthode permet une meilleure estimation du spectre que les méthodes usuelles utilisant les “pseudo- $C_\ell$ ”.

Nous avons aussi réalisé des simulations d’observations par plusieurs expériences simultanées. Ces simulations reproduisent les zones observées, les résolutions angulaires et les niveaux de bruit de 6 cartes parmi les plus récemment mesurées, correspondant à

- a) 3 canaux d’observation (appelés Q, V et W) de l’expérience WMAP
- b) 2 zones d’observations (imbriquées) de l’expérience BOOMERANG
- c) 3 zones d’observations de l’expérience ACBAR.

La possibilité de disposer d’un estimateur unique (pour chaque bande de fréquence) agrégeant toutes les expériences est un avantage important de notre méthode. Nous illustrons sur les simulations la façon dont les différentes expériences se combinent d’un point de vue statistique, en appliquant notre



F . 4.9 – Dans chaque bande de fréquences, corrélation empirique de l’estimateur utilisant l’ensemble des données, avec chacun des estimateurs utilisant des jeux partiels.

estimateur à des jeux partiels de données – successivement a) seul, b) seul et c) seul – puis à l’ensemble des données des trois expériences. La figure 4.9 montre la corrélation entre les estimateurs issus de chaque jeu partiel avec l’estimateur utilisant l’ensemble des données. La figure 4.10 représente l’erreur quadratique moyenne de chacun des estimateurs. Cela permet de quantifier l’intuition que :

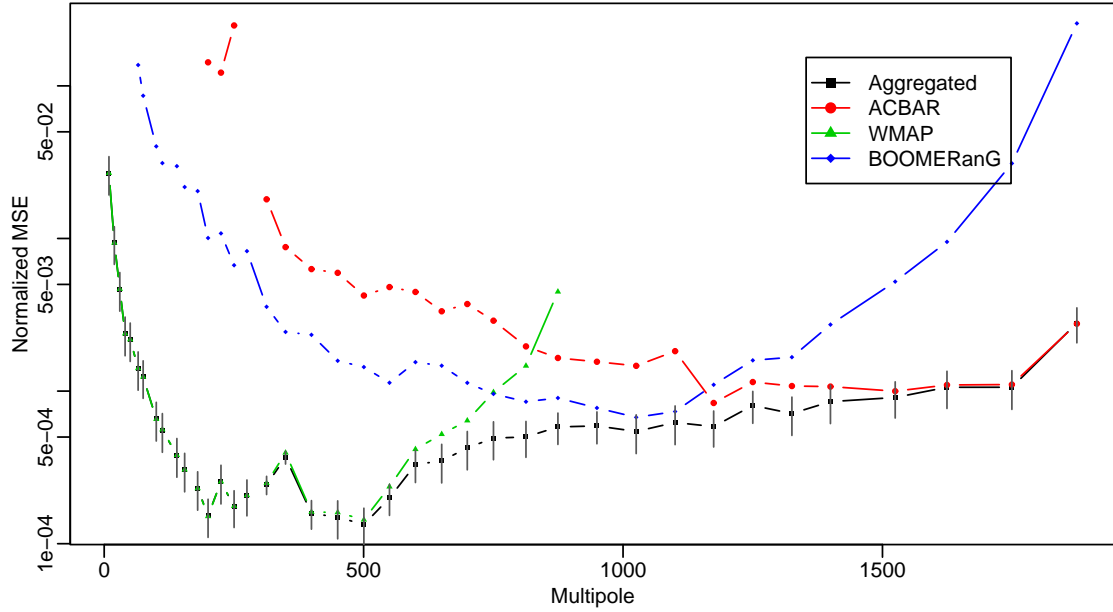
- l’information statistique est principalement apportée par des expériences différentes selon les fréquences : WMAP aux basses fréquences, ACBAR aux très hautes fréquence, etc.,
- le gain de notre approche est surtout important aux fréquences pour lesquelles plusieurs expériences apportent une quantité comparable d’information.

### Séparation de sources : extraction du Fond diffus

Nous faisons une autre utilisation du caractère doublement localisé des needlets, pour la séparation de sources, c’est-à-dire que nous réalisons le programme de Tegmark (cité dans le Chapitre 3, page 58). L’objectif est de reconstruire, à partir des données qui sont des superpositions du CMB et des autres composantes astrophysiques, une carte de CMB “pur” par combinaison de cartes de mélanges à différentes fréquences lumineuses (voir le Chapitre 3, équations (3.22)–(3.21)–(3.23)). La qualité d’une carte se mesure au fait qu’elle soit complète (sur toute la sphère), d’erreur (MSE) minimale et de résolution angulaire maximale.

Nous nous plaçons dans le cadre du modèle décrit par l’équation (3.14). L’estimateur (3.22)–(3.21)–(3.23) est mis en oeuvre dans le domaine des coefficients de needlets : une fois calculés les coefficients  $\beta_{j,k} = [\beta_{j,k,1}, \dots, \beta_{j,k,N_{freq}^{(j)}}]^T$ , où  $\beta_{j,k,v}$  est le coefficients de l’observation à la  $v^{\text{ième}}$  fréquence lumineuse, la carte de CMB reconstruite est

$$\widehat{X} := \sum_{j=0}^J \sum_{k=1}^{N_j} \widehat{\beta}_{j,k} \widetilde{\psi}_{j,k}$$



F . 4.10 – Moyennes empiriques des quantités  $\frac{(\widehat{C}_j - C^{(j)})^2}{(C^{(j)})^2}$  pour les estimateurs utilisant l'ensemble des données ("Aggregated", pour lequel des barres d'erreur sont ajoutées, de demi-longueur valant deux écarts-types empiriques) et pour les estimateurs utilisant des jeu partiels.

où

$$\widehat{\beta}_{j,k} := (\mathbf{1}^T \widehat{R}_{j,k}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \widehat{R}_{j,k}^{-1} \beta_{j,k} .$$

La matrice  $\widehat{R}_{j,k}$  estime la matrice de covariance de  $\beta_{j,k}$  par

$$\widehat{R}_{j,k} := (\text{Card } V(k))^{-1} \sum_{p \in V(k)} \beta_{j,p} \beta_{j,p}^T$$

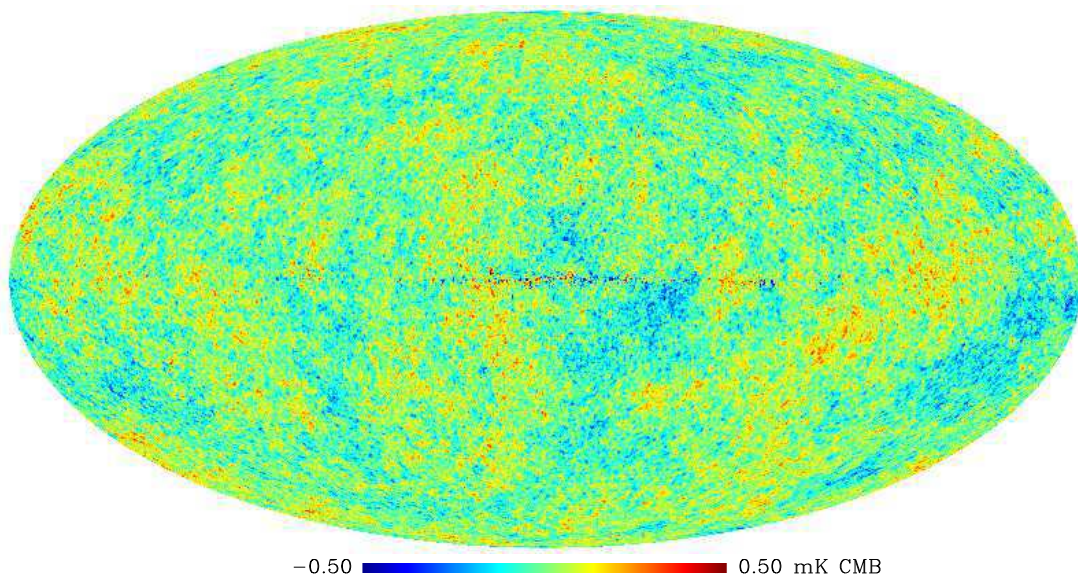
où  $V(k)$  est un voisinage de  $k$ , ce qui suppose implicitement une certaine régularité du niveau de bruit, de sorte que le bruit puisse être considéré comme localement homogène.

L'avantage de cette méthode est qu'elle permet de s'adapter au niveau local du bruit et de la contamination par les avant-plans (sources autres que le CMB), tout en tenant compte également des puissances relatives, selon la fréquence, du signal et du bruit. Elle permet également d'utiliser à chaque échelle  $j$  l'ensemble des cartes observées à cette échelle, ce qui n'est pas le cas des méthodes utilisant les cartes dans le domaine direct. Cette capacité est utile car les différentes cartes d'entrée ne sont pas toutes à la même résolution.

Des simulations ont permis de valider cette approche et de la calibrer. Son application aux données réelles constituées des cinq cartes délivrées par l'expérience WMAP après 5 ans d'observations nous permet d'obtenir sans doute la carte complète de CMB la plus précise à ce jour (figure 4.11).



5 year needlet ILC map



F . 4.11 – Carte de CMB reconstruite à partir des données WMAP.

## 4.4 Note technique

Du point de vue informatique, nos méthodes ont été développées principalement avec le logiciel octave<sup>1</sup>. Nous avons pour cela interfacé sous octave des logiciels spécifiques d'analyse sur la sphère : HEALPix<sup>2</sup> et GLESP<sup>3</sup> (parmi les autres logiciels d'analyse sur la sphère, mentionnons SHTools<sup>4</sup>). Certaines fonctions nécessitant un grand nombre d'itérations ont été compilées en C. Une partie des fonctions Octave développées constituent le logiciel SphereLab<sup>5</sup>.

La plupart des procédures implémentées (optimisation du critère géométrique pour la localisation des needlets ; analyse et synthèse en needlets ; estimation spectrale ; etc.) sont suffisamment rapides pour tourner sur un ordinateur individuel jusqu'à des résolutions de l'ordre d'un million de pixels<sup>6</sup> (en quelques secondes ou quelques minutes) voire d'une dizaine de millions de pixels (en quelques heures pour une analyse complète sur un processeur de 2×2 GHz et 2 Go de mémoire). Néanmoins, pour les simulations massives sur des cartes à hautes résolutions, comme celles de la Section 4.2, nous avons utilisé le cluster de 88×2.3 GHz et 176 Go de mémoire distribuée-partagée de l'équipe Adamis.

## 4.5 Perspectives

Des perspectives nombreuses et variées s'inscrivent dans la continuation de ce travail. Notre méthode d'estimation spectrale à partir des coefficients de needlet a été testée sur des simulations intégrant une grande partie de la complexité du modèle physique. La prochaine étape est naturellement de l'appliquer aux données disponibles issues des observations. Il sera possible d'utiliser simultanément l'ensemble des données publiques, ce qui devrait permettre de préciser les valeurs du spectre de puissance actuellement estimées. En dehors d'un important travail technique de préparation des données, cette application nécessitera d'un point de vue méthodologique la prise en compte des multiples petites zones de données corrompues qui criblent le ciel (voir la figure 1.12). Ces zones (qui correspondent physiquement à des sources ponctuelles d'émission lumineuse) pourraient être traitées exactement comme nous l'avons fait dans le cas de la séparation de sources (chapitre 8).

Concernant toujours les applications, l'utilisation de needlets optimalement localisées est envisageable pour d'autres analyses statistiques du CMB : séparation des différentes sources (autres que le CMB), tests de gaussianité et d'isotropie. Leur utilisation dans d'autres domaines faisant intervenir l'analyse de champs sphériques (comme la géophysique) pourrait aussi s'avérer fructueuse.

Du point de vue théorique, les résultats obtenus demandent à être approfondis et généralisés : au-delà de la convergence, il est nécessaire – et probablement peu coûteux en s'appuyant les résultats de Baldi et al. [151] – d'obtenir distributions asymptotiques et vitesses de convergence (toujours au sens inhabituel de la définition 4.1) des estimateurs proposés. Il serait intéressant en outre de les comparer sous cet angle aux autres estimateurs utilisés en pratique (par exemple ceux présentés dans la section 3.2), et d'étudier leur efficacité asymptotique.

---

<sup>1</sup>Logiciel libre semblable à Matlab. <http://www.gnu.org/software/octave>

<sup>2</sup>Hierarchical Equal Area isoLatitude Pixelization. <http://healpix.jpl.nasa.gov>

<sup>3</sup>Gauss-Legendre Sky Pixelization. <http://www.glesp.nbi.dk>

<sup>4</sup>Spherical Harmonics Tools. <http://www.ipgp.jussieu.fr/~wieczor/SHTOOLS/SHTOOLS.html>

<sup>5</sup>[http://www.apc.univ-paris7.fr/APC\\_CS/Recherche/Adamis/Outreach/spherelab-fr.php](http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/Outreach/spherelab-fr.php)

<sup>6</sup>nside  $\leq 512$  avec HEALPix, permettant le calcul des coefficients de Fourier pour  $\ell \leq 1500$

Jusqu'à présent, on a supposé le bruit indépendant de point à point, et le niveau de bruit  $\sigma(\xi)$  connu. Une généralisation intéressante serait de s'affranchir de ces hypothèses : modéliser la dépendance du bruit dans un modèle discret ou continu, et en estimer le niveau à partir des observations.

S'agissant des recherches à venir dans le domaine de l'analyse sur la sphère, deux aspects paraissent importants. Le premier est une comparaison plus systématique des différentes constructions d'ondelettes sphériques. Cela permettrait peut-être de trouver des conditions plus générales assurant les qualités de frame et de localisation. Par exemple, Geller et Mayeli [21] ont récemment construit des needlets "presque ajustées" (au sens des frames) à décroissance exponentielle. Le second est la généralisation des needlets à des champs aléatoires sphériques de matrices symétriques (au lieu de fonctions ou champs aléatoires à valeurs réelles). De telles champs matriciels correspondent en effet à l'observation de la polarisation du Fond diffus cosmologique, qui est l'une des principales avancées observationnelles attendues dans les prochaines années.

## **Deuxième partie**

### **Articles de revues**

## Chapitre 5

# Practical wavelet design on the sphere

Ce chapitre reproduit l'article du même nom, F. Guilloux, G. Faÿ, J.-F. Cardoso, *Applied and Computational Harmonic Analysis* (à paraître)\*. Seules certaines notations ont été modifiées pour la cohérence du manuscrit.

---

**Abstract.** This paper considers the design of isotropic analysis functions on the sphere which are perfectly limited in the spectral domain and optimally localized in the spatial domain. This is motivated by the need of localized analysis tools in domains where the data is lying on the sphere, *e.g.* the science of the Cosmic Microwave Background. Our construction is derived from the localized frames introduced by Narcowich, Petrushev et Ward [33]. The analysis frames are optimized for given applications and compared numerically using various criteria.

### Sommaire

---

Introduction . . . . .	<b>81</b>
5.1 Needlets frames . . . . .	<b>82</b>
5.1.1 Notations and background . . . . .	82
5.1.2 Needlet frames . . . . .	83
5.1.3 Our setting . . . . .	86
5.1.4 Practical computation of needlet coefficients . . . . .	87
5.2 Design of optimally localized wavelets . . . . .	<b>87</b>
5.2.1 Examples. . . . .	88
5.2.2 $\mathbb{L}^2$ -concentration and variations . . . . .	90
5.2.3 Statistical criterion for optimal analysis with missing data . . . . .	93
5.3 Examples, numerical results . . . . .	<b>96</b>
5.3.1 Comparison of filters for various criteria . . . . .	96
5.3.2 Robustness of needlets coefficients . . . . .	97
5.3.3 Some MISE-optimal filters for axisymmetric weight functions . . . . .	97
5.4 Conclusion . . . . .	<b>100</b>
5.5 Proofs . . . . .	<b>101</b>

---

---

\*<http://dx.doi.org/10.1016/j.acha.2008.03.003>

## Introduction

Data defined on the sphere are found in numerous application domains. During the past decade, *localized* analysis for spherical data has motivated many developments, whether it be in cosmology [28, 133], geophysics [184, 51, 15, and the references therein], biomedical engineering [see *e.g.* 191, and the references therein], computer vision, hydrodynamics, etc. When the data are available on the whole sphere, they can be analyzed by projection onto the spherical harmonics, with the spherical harmonic transform (SHT) providing the spherical equivalent of the Fourier Series on the circle. But in many practical situations, data are defined or available only on a fraction of the sphere. For example, cosmologists try to give sharp estimates of the cosmic microwave background (CMB) or its power spectrum but strong *localized* foreground emissions superimpose to the CMB making it unreliable to use the full sphere for accurate spectral estimation. Even when data are available over the whole sphere, spatially localized tools are invaluable to process non stationary fields or stationary fields corrupted by non-stationary noise. In such situations, spherical harmonics are not adequate due to their lack of spatial localization. A mathematically elegant solution of proven practical efficiency is provided by multiscale and wavelet theories. These are well developed for functions defined on Euclidean spaces ; some practical aspects of their extension to the sphere is the focus of this paper.

Adaptation to the sphere of the “wavelet” transform (in the broad sense of filtering by spatially and spectrally localized functions) was introduced a dozen years ago [39, 46, 7, 34, 38, 20]. Since then, Antoine and Vandergheynst [5] showed that any Continuous Wavelet Transform (CWT) on the sphere can be viewed locally as a regular CWT on the tangent plane, thanks to the stereographic correspondence between the sphere and the plane [5, 50]. A discretized version of this CWT has been presented by [6], leading to wavelet frames. This approach has already been followed in astrophysics for the analysis of the CMB [see *e.g.* 142, 28]. The stereographic projection on the tangent plane helps to identify the similitude group on the sphere and derive the CWT in a group-theoretical manner but in practice, the whole analysis is performed directly on the sphere. In the above-mentioned references, the spherical wavelets that are given as examples or are used in the real data analysis are defined on the tangent plane and transformed (once for all) to the sphere by the stereographic projection. For instance, the Mexican hat wavelet family is constructed from the Gaussian density function on the  $\mathbb{R}^2$ -plane ; there is no deep rationale for using it as a spherical function. Moreover, these wavelets are defined in the spatial domain and have infinite support in the frequency domain (which must be truncated in practice).

In the present work, we follow and extend the approach of Narcowich, Petrushev and Ward [32, 33] and their construction of “needlets”. A similar construction can be found in Starck et al. [134]. The needlet transform has important distinctive properties. Firstly it is intrinsically spherical. No intermediate tangent plane is needed to define it. Secondly, it does not depend on the particular spherical pixelization (sampling) chosen to describe the data. Thirdly, although the needlets still have an excellent spatial localization, they have a finite spectral support adjustable at will. They are axisymmetric (which is convenient when dealing with statistically isotropic random fields) and thus the needlet coefficients are easily computed in the spherical harmonic (Fourier) domain. Filtering is obtained by multiplication of the spherical harmonic coefficients by well chosen window functions (which is equivalent to a convolution in spatial domain). Needlets are well defined in theory and the statistical properties of their coefficients have already been established for isotropic Gaussian fields by Baldi et al. [151]. However, the performance of a needlet-based analysis depends on the particular shape of the needlet.

We are concerned with the development of a flexible spectral analysis on the sphere which remains practical at high resolution. For instance the Planck space mission<sup>1</sup> of the European Spatial Agency will produce CMB maps as large as 50 mega-pixel with reliable multipoles up to order  $\ell \simeq 4000$ .

This paper focuses on the design issue, namely the optimization of the window functions which define the needlet transform. We consider only band-limited needlets. This choice is motivated by applications in high-precision cosmology. Indeed, the angular power spectrum of the CMB varies over a large dynamic range (power-law decay) so that using band-limited wavelets guarantees the absence of spectral leakage from the very large low-frequency components to the much weaker high-frequency components.

Strict simultaneous confinement of a needlet in space and frequency is, of course, impossible. One of the most famous expression of this fact is the Heisenberg uncertainty principle which has been extended to the sphere [34, 22]. Recently Fernandez [13] constructed band-limited wavelets which minimize the uncertainty product. However, paraphrasing Slepian [42], this criterion which is relevant to quantum mechanics may not be as meaningful as a simple  $\mathbb{L}^p$  criterion in our context.

In this contribution, we investigate the design of band-limited spectral window functions in two different directions : 1) By requesting the best spatial localization of associated needlets, in an energy-sense ( $\mathbb{L}^2$ ) which is easily solved. This is an application of the work of [41] which adapted to the sphere the problem solved by Slepian [187] on the real line, giving rise to the well known prolate spheroidal wave functions (PSWF). 2) By the following statistical problem : given some region of the sphere in which the data is missing (or has to be discarded), we seek to minimize, in the needlet analysis, the error introduced by the missing data. More criteria and applications to cosmological science will be given in a future work.

The paper is organised as follows. In Section 5.1, we expose the general construction of needlets. In Section 5.2, we define and optimize the two above-mentioned criteria (geometrical and statistical), yielding needlets which are optimal in some well-defined sense. Their efficiency is illustrated in Section 5.3 with numerical simulations following the model of a masked observation of the CMB. Conclusions are given in Section 5.4 and proofs are postponed to Section 5.5.

## 5.1 Needlets frames

### 5.1.1 Notations and background

Denote  $\mathbb{S}^2$  the unit sphere in  $\mathbb{R}^3$  with generic element  $\xi = (\theta, \varphi)$  in spherical polar coordinates :  $\theta \in [0, \pi]$  is the colatitude and  $\varphi \in [0, 2\pi[$  the longitude. Let  $\mathbb{L}^2(\mathbb{S}^2)$  be the space of complex-valued square integrable functions on  $\mathbb{S}^2$  under the Lebesgue measure  $d\xi = \sin \theta d\theta d\varphi$ . Endowed with the inner product  $\langle f, g \rangle := \int_{\mathbb{S}^2} f(\xi) g^*(\xi) d\xi$ ,  $\mathbb{L}^2(\mathbb{S}^2)$  is a Hilbert space. Let  $\|\cdot\|$  denote the associated norm on  $\mathbb{L}^2(\mathbb{S}^2)$ . The usual complex spherical harmonics on  $\mathbb{S}^2$  are denoted  $(Y_{\ell, m})_{\ell \geq 0, -\ell \leq m \leq \ell}$ . They form an orthonormal basis of  $\mathbb{L}^2(\mathbb{S}^2)$ . The Legendre polynomials  $L_\ell$  are normalized according to the condition

---

<sup>1</sup>see [www.rssd.esa.int/Planck/](http://www.rssd.esa.int/Planck/)

$L_\ell(1) = \frac{2\ell+1}{4\pi}$  (or  $\int_{-1}^{+1} L_\ell(z)^2 dz = \frac{2\ell+1}{8\pi^2}$ ), which leads to coefficient-free properties like

$$L_\ell(\xi' \cdot \xi) = \sum_{m=-\ell}^{\ell} Y_{\ell m}^*(\xi) Y_{\ell m}(\xi')$$

(the summation formula) and

$$\int_{\mathbb{S}^2} L_\ell(\eta \cdot \xi) L_{\ell'}(\eta' \cdot \xi) d\xi = \delta_{\ell\ell'} L_\ell(\eta \cdot \eta'). \quad (5.1)$$

In the following, we consider a random field  $X \in \mathbb{L}^2(\mathbb{S}^2)$ . Its spherical harmonic coefficients or multipole moments are denoted  $a_{\ell,m} = \langle X, Y_{\ell,m} \rangle$ .  $\mathbb{L}^2(\mathbb{S}^2)$  can be decomposed into orthogonal harmonic subspaces :  $\mathbb{L}^2(\mathbb{S}^2) = \bigoplus_{\ell \geq 0}^{\perp} \mathbb{H}_\ell$ , where  $\mathbb{H}_\ell$  is the linear span of  $Y_{\ell m}, m = -\ell, \dots, \ell$ . The number  $\ell$  is referred to as the multipole number or frequency (understood as a spatial frequency). Let  $\Pi_\ell$  be the orthogonal projection on  $\mathbb{H}_\ell$ . It writes

$$\Pi_\ell X(\xi) = \sum_{m=-\ell}^{\ell} \langle X, Y_{\ell,m} \rangle Y_{\ell,m}(\xi) = \int_{\mathbb{S}^2} X(\xi') L_\ell(\xi \cdot \xi') d\xi', \quad (5.2)$$

$\xi \cdot \xi' = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\varphi - \varphi')$  being the usual dot product on  $\mathbb{S}^2$ .

A pattern on  $\mathbb{S}^2$  which depends on the colatitude  $\theta$  only is said to be axisymmetric. The convolution of a bounded axisymmetric function  $H(\xi) = h(\cos \theta)$  with an arbitrary spherical function  $X$  is well defined through

$$H * X(\xi) = \int_{\mathbb{S}^2} h(\xi \cdot \xi') X(\xi') d\xi' \quad (5.3)$$

The convolution theorem holds :

$$H * X = \sum_{\ell,m} h_\ell a_{\ell,m} Y_{\ell,m}. \quad (5.4)$$

where  $a_{\ell,m} = \langle X, Y_{\ell,m} \rangle$  are the multipole moments of  $X$  and  $h_\ell$  are the Legendre series coefficients of  $h$ , i.e.  $h = \sum_{\ell \in \mathbb{N}} h_\ell L_\ell$ . An isotropic wavelet analysis can be computed either in the spatial (i.e. direct) domain using (5.3) or in the spherical harmonic domain using (5.4). In practice, the latter, which amounts to multiply the multipole moments of the field of interest  $X$  by a spectral window ( $h_\ell$ ), is trivially implemented as soon as a SHT is available.

### 5.1.2 Needlet frames

This section reviews the construction and first properties of the *needlets* presented in [33, 32] (see also Baldi et al. [151]).

**Definition 5.1** (Quadrature). *Let  $K$  be a finite index set and  $\{\xi_k\}_{k \in K} \in (\mathbb{S}^2)^{|K|}$  a set of distincts quadrature points on the sphere, associated with a set  $\{\lambda_k\}_{k \in K} \in \mathbb{R}^{|K|}$  of quadrature weights.  $\{(\xi_k, \lambda_k)\}_{k \in K}$  is said to provide an exact Gauss quadrature formula at degree  $\ell^{\max}$  if*

$$\forall X \in \bigoplus_{\ell=0}^{\ell^{\max}} \mathbb{H}_\ell, \quad \int_{\mathbb{S}^2} X(\xi) d\xi = \sum_{k \in K} \lambda_k X(\xi_k).$$

*This quadrature formula is said to be positive-weight if  $\lambda_k > 0, k \in K$ .*



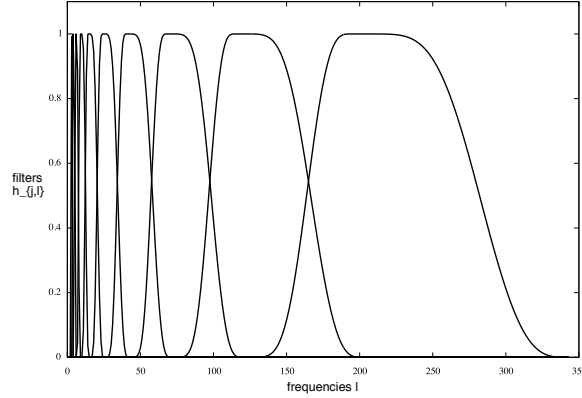


FIG. 5.1 – First 10 windows defined by (5.5),  $\mathbf{h}$  being a spline of order 7 compactly supported on  $[\frac{1}{B}, B]$  with  $B = 1.7$ .

**Remark 5.1.** We refer to Doroshkevich et al. [73] for an example of a proper choice of quadrature points and weights —GLESP— that satisfy this property. Other pixelization schemes, such as the one implemented in HEALPIX (Gorski et al. [84]), provide an approximate quadrature with a number of data points of order  $C(\ell^{max})^2$  and a number of quadrature weights of order  $\frac{1}{C(\ell^{max})^2}$  for some positive constant  $C$ .

Assume that  $\mathbf{h}$  is a non-negative function supported on  $[\frac{1}{B}, B]$  and verifying  $\mathbf{h}(x) + \mathbf{h}(B^{-1}x) = 1, x \in [1, B]$ , for some  $B > 1$ . For any non-negative  $j, \ell$ , put

$$h_{j,\ell} := \mathbf{h}(B^{-j}\ell) \quad , \quad b_{j,\ell} := \sqrt{h_{j,\ell}} \quad (5.5)$$

and  $\ell_j^{max} := \max\{\ell : h_{j,\ell} \neq 0\}$ . Note that  $\ell_j^{max} \leq B^{j+1}$ . This is illustrated in Figure 5.1. In the following,  $j$  is referred to as the *scale*, in analogy with the multiresolution terminology.

Assuming a pixelization  $\{\xi_{j,k}, \lambda_{j,k}\}_{k \in K_j}$  attached to each scale  $j$ , one defines :

**Definition 5.2** (Needlets and Needlet coefficients). *For every  $j \in \mathbb{N}$  and every index  $k \in K_j$  the function*

$$\psi_{j,k}(\xi) = \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\ell_j^{max}} b_{j,\ell} L_\ell(\xi \cdot \xi_{j,k}), \quad (5.6)$$

*is called a needlet. For  $X \in \mathbb{L}^2(\mathbb{S}^2)$ , the inner products  $\langle X, \psi_{j,k} \rangle$  are called needlet coefficients and are denoted  $\beta_{j,k}$ .*

Up to a rotation of the sphere putting  $\xi_{j,k}$  on the North pole and to the multiplicative term  $\sqrt{\lambda_{j,k}}$ , all the needlets of a given scale  $j$  have exactly the same shape. In particular, they are axisymmetric. When  $\mathbf{b} := \sqrt{\mathbf{h}}$  is  $C^\infty$ , the needlet  $\psi_{j,k}$  is almost exponentially localized around  $\xi_{j,k}$ , see Eq (5.12) below. This is a major quality of those space frequency kernels. The following Proposition states that the decomposition of  $\mathbb{L}^2(\mathbb{S}^2)$  on the needlets family built with (5.5) can be viewed as some harmonic smoothing operation (see the illustration on Figure 5.2), and that this family is a tight frame.

**Proposition 5.1.** *Let  $j \in \mathbb{N}$ . Assume that  $\{(\xi_{j,k}, \lambda_{j,k})\}_{k \in K_j}$  provides an exact and positive-weight quadrature formula at degree  $2\ell_j^{\max}$ . Then*

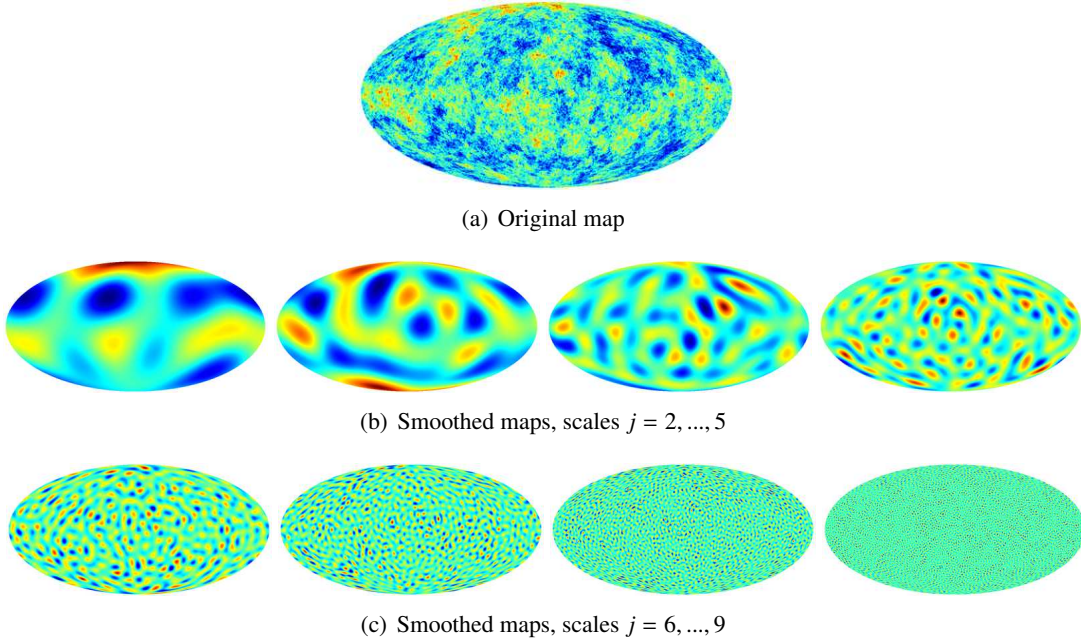
$$\sum_{k \in K_j} \beta_{j,k} \psi_{j,k} = \Phi_j X$$

with

$$\Phi_j := \sum_{\ell \in \mathbb{N}} h_{j,\ell} \Pi_\ell.$$

Assume that for any  $j \in \mathbb{N}$ ,  $\{(\xi_{j,k}, \lambda_{j,k})\}_{k \in K_j}$  provides an exact and positive-weight quadrature formula at degree  $2\ell_j^{\max}$ . Then,

$$\forall X \in \mathbb{L}^2(\mathbb{S}^2), \quad X = \sum_{j \in \mathbb{N}} \sum_{k \in K_j} \beta_{j,k} \psi_{j,k} \quad \text{and} \quad \|X\|^2 = \sum_{j \in \mathbb{N}} \sum_{k \in K_j} |\beta_{j,k}|^2.$$



F . 5.2 – Input map of a CMB sky (from WMAP), and corresponding smoothed maps (with the spline filters of Figure 5.1).

**Remark on Terminology** The analysis of an input field  $X$  in the way described above is called *filtering*. This filtering has two equivalent expressions, in the spatial and in the spectral domains ; see the convolution formula (5.4). These expressions involves two “dual” mathematical objects : the functions  $h_{j,\cdot}$  and  $b_{j,\cdot}$  of the frequency  $\ell$ , called *window functions* or *spectral windows*, and the spherical functions  $\psi_{j,k}$  called *needlets*, which are nothing else but the rotated axisymmetric functions built from the Legendre transform of  $b_{j,\cdot}$ . (see Definition 5.2). We call *filter* either of the two above objects, when the domain (spatial or spectral) is not specified.

### 5.1.3 Our setting

Narcowich *et al.* [32] extended the previous approach to build dual wavelet frames in a dyadic context. We generalize this construction, get rid of the  $\mathcal{B}$ -adic scheme, and construct needlets with arbitrary frequency support. The scales are now indexed by  $j \in \mathcal{J}$  a countable set. We shall design strictly band-limited needlets with support  $I_j := [\ell_j^{\min}, \ell_j^{\max}]$ ,  $\ell_j^{\min} > 0$  if  $j \geq 0$ . This design will be performed in the spherical harmonic domain.

Since the needlet coefficients  $\beta_{j,k}$  and  $\beta_{j',k'}$  of a Gaussian stationary (thus isotropic) field are independent if  $I_j \cap I_{j'} = \emptyset$ , it is advantageous to choose the bands  $I_j$  to overlap as little as possible. Other choices are possible ; for instance Starck *et al.* [134] take overlapping spectral windows supported on  $[0, 2^j]$ .

From any windows family  $(b_j)_{j \in \mathcal{J}}$  such that  $\forall \ell \in \mathbb{N}, \sum_{j \in \mathcal{J}} (b_{j,\ell})^2 > 0$ , define the synthesis windows  $\tilde{b}_j$  by

$$\forall j \in \mathcal{J}, \forall \ell \in \mathbb{N}, \tilde{b}_{j,\ell} = \frac{b_{j,\ell}}{\sum_{j' \in \mathcal{J}} (b_{j',\ell})^2} \quad (5.7)$$

and put  $h_j := \tilde{b}_j b_j$  so that  $\forall \ell, \sum_{j \in \mathcal{J}} h_{j,\ell} = 1$  follows. We retain Definition 5.2 for the *decomposition* needlets and needlets coefficients and further define the *reconstruction* needlets as

$$\tilde{\psi}_{j,k}(\xi) = \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\ell_j^{\max}} \tilde{b}_{j,\ell} L_\ell(\xi \cdot \xi_{j,k}) . \quad (5.8)$$

**Proposition 5.2.** *Assume that there exists positive constants  $A, B$  such that*

$$\forall \ell \in \mathbb{N}, A \leq \sum_{j \in \mathcal{J}} |b_{j,\ell}|^2 \leq B . \quad (5.9)$$

*Assume that for any  $j \in \mathcal{J}$ , the set  $\{(\xi_{j,k}, \lambda_{j,k})\}_{k \in K_j}$  provides an exact and positive-weight quadrature formula at degree  $2\ell_j^{\max}$ . Then, the family  $\{\psi_{j,k}\}$  is a frame with frame bounds constant  $A$  and  $B$ . Its dual frame is the family  $\{\tilde{\psi}_{j,k}\}$ . In particular*

$$\forall X \in \mathbb{L}^2(\mathbb{S}^2), X \stackrel{\mathbb{L}^2(\mathbb{S}^2)}{=} \sum_{j \in \mathcal{J}} \sum_{k \in K_j} \beta_{j,k} \tilde{\psi}_{j,k} \text{ and } \|X\|^2 = \sum_{j \in \mathcal{J}} \sum_{k \in K_j} \tilde{\beta}_{j,k} \beta_{j,k} , \quad (5.10)$$

with  $\tilde{\beta}_{j,k} := \langle X, \tilde{\psi}_{j,k} \rangle$ .

Define the analysis, synthesis and smoothing operators at scale  $j \in \mathcal{J}$  by  $\Psi_j = \sum_\ell b_{j,\ell} \Pi_\ell$ ,  $\tilde{\Psi}_j = \sum_\ell \tilde{b}_{j,\ell} \Pi_\ell$  and  $\Phi_j = \tilde{\Psi}_j \Psi_j$ , respectively. Then, the exact reconstruction formula  $\sum_j \Phi_j = \text{Id}$  holds true.

An example of analysis/synthesis window functions following this scheme is shown in Figure 5.3, for which we take optimally concentrated PSWF (see Section 5.2) functions for analysis. It illustrates the fact that this choice does not lead to well localized synthesis needlets (as their spectral shapes are non smooth). This has a sensible impact on reconstruction procedures (such as bandwise filtering). However, this may not be a shortcoming if, for instance, one is concerned only with the statistical properties of the needlet coefficients  $\beta_{j,k}$ .

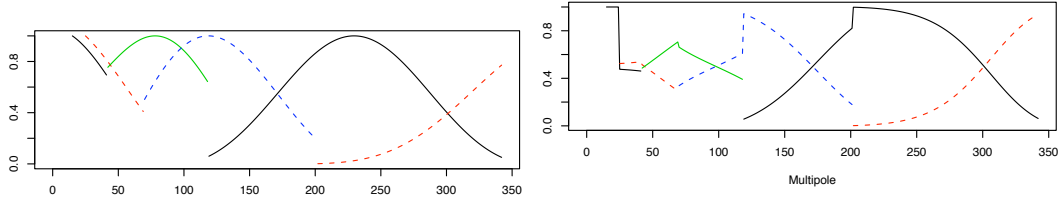


FIG. 5.3 – PSWF analysis window functions (top) and dual synthesis window functions (bottom).  $\mathcal{B}$ -adic scheme with  $B = 1.7$ ,  $j = 6, \dots, 11$ .

### 5.1.4 Practical computation of needlet coefficients

Although theoretically well-defined, the evaluation of inner products  $\langle X, \psi_{j,k} \rangle$  in the direct space is practically unfeasible at high resolutions. These coefficients are much more easily computed via direct and inverse spherical harmonic transforms. The following Proposition gives a way to compute the coefficients as samples of spherical maps.

**Proposition 5.3.** *The needlet coefficients verify  $\beta_{j,k} = \sqrt{\lambda_{j,k}} \Psi_j X(\xi_{j,k})$ .*

The computation of the smoothed field  $\Psi_j X$  is performed in the spherical harmonic domain by multiplying the multipole coefficients  $a_{\ell,m}$  of  $X$  by the factors  $b_{j,\ell}$  and the needlet coefficients  $\beta_{j,k}$  are retrieved as the values of  $\Psi_j X$  at the quadrature points  $\xi_{j,k}$  up to a multiplicative term. Starting from the field  $X$  sampled at some quadrature points, this is summed up by the following diagram

$$\{X(\xi_k)\}_{k \in K} \xrightarrow{\text{SHT}} \{a_{\ell,m}\}_{\ell,m} \xrightarrow{\times} \{b_{j,\ell} a_{\ell,m}\}_{\ell,m} \xrightarrow{\text{SHT}^{-1}} \{(\lambda_{j,k})^{-1/2} \beta_{j,k}\}_{k \in K_j} \quad (5.11)$$

whereas the synthesis operation is summed up by

$$\{(\lambda_{j,k})^{-1/2} \beta_{j,k}\}_{k \in K_j} \xrightarrow{\text{SHT}} \{b_{j,\ell} a_{\ell,m}\}_{\ell,m} \xrightarrow{\times} \{\tilde{b}_{j,\ell} b_{j,\ell} a_{\ell,m}\}_{\ell,m} \xrightarrow{\text{SHT}^{-1}} \{\Phi_j X(\xi_{j,k})\}_{k \in K_j}$$

Software packages for spherical pixelization, such as HEALP<sup>2</sup>, GLESP or SHTOOLS<sup>2</sup>, include efficient implementations of the direct and inverse spherical harmonic transforms. This makes the computation easy and tractable even at high resolution. The needlet coefficients at a given scale  $j$  can be visualized as a pixelized map. If the quadrature weights  $\{\lambda_{j,k}\}$  are equal, the smoothed maps of Fig. 5.2, which are the outputs of the processing (5.11), provide a precise and easily interpretable picture of the space-frequency analysis.

**Remark 5.2.** *The quadrature points and weights  $\{(\xi_{j,k}, \lambda_{j,k})\}_{k \in K_j}$  use to define the needlet coefficients  $\beta_{j,k}$  and to sample the smoothed field  $\Phi_j X$  may be chosen identical to  $\{(\xi_k, \lambda_k)\}_{k \in K}$  used to sample the input field  $X$ . However, for data compression and computational efficiency, it is also possible to use, at each scale  $j$ , the smallest possible  $K_j$  for which gives the exact quadrature up to order  $2\ell_j^{\max}$ .*

## 5.2 Design of optimally localized wavelets

In this section, we define some criteria to compare the window profiles. Some of them are easily optimized, others are only investigated numerically. We first give some examples of generic needlet

<sup>2</sup>available from <http://www.ipgp.jussieu.fr/~wieczor/SHTOOLS/SHTOOLS.html>

profiles (Section 5.2.1). Then, we restrict ourselves to a single scale  $j$  and an associated band  $L := [\ell^{min}, \ell^{max}]$ . The superscript  $(j)$  will be omitted in the notations when no confusion is possible. We present the  $\mathbb{L}^2$  (Section 5.2.2) and statistical (Section 5.2.3) criteria, with practical implementation details on their optimization.

### 5.2.1 Examples.

Narcowich *et al.* [33] have derived the following theoretical bound that controls the decay of the needlets. In the  $\mathcal{B}$ -adic case, if  $\mathbf{b} := \sqrt{\mathbf{h}}$  is  $M$ -times continuously differentiable

$$|\psi_{j,k}(\xi)| \leq \frac{C B^{j-1}}{1 + (B^{j-1} \arccos(\xi \cdot \xi_{j,k}))^M} \quad (5.12)$$

for some constant  $C = C(\mathbf{b})$ . This condition still allows a wide range of possibilities for designing the function  $\mathbf{b}$ . Without restricting ourselves to the  $\mathcal{B}$ -adic case, we shall describe practical methods for the effective optimization of a spectral window  $b_j$  based on non-asymptotic, application-dependent criteria.

In order to illustrate some points of interest, we compare in Figure 5.4 the azimuthal profiles of various axisymmetric needlets. More precisely, we consider the shape of the needlets defined by (5.6) with  $\xi_k = (0, 0)$ ,  $\lambda_k = 1$ . Being centered at the North pole, they are functions of  $\theta$  only. We consider only the 9<sup>th</sup> dyadic scale :  $I = [256, 1024]$  and compare heuristically five (families of) window functions. Note that the last two windows are not limited to band  $L$ .

1. Square roots of splines of various orders. For any odd integer  $M$ , there exists a spline function  $\mathbf{h}$  of order  $M$ , non-negative, compactly supported on  $[\frac{1}{2}, 2]$  and verifying  $\mathbf{h}(x) + \mathbf{h}(x/2) = 1$ ,  $x \in [1, 2]$ . The  $b_{j,\ell}$  are defined by (5.5).
2. Best concentrated Slepian functions in caps of various radii (cf Section 5.2.2). The window function  $b_{j,\ell}$  is the minimizer of the criterion (5.18). It is band-limited on  $L$  and optimally concentrated in a polar cap  $\Omega_{\theta_0} = \{\xi : \theta \leq \theta_0\}$ ,  $\theta_0$  being a free parameter.
3. Denoting  $G$  a primitive of the  $C^\infty$  function  $g : x \mapsto e^{-\frac{1}{1-x^2}} 1_{(-1,1)}(x)$ , set

$$\mathbf{b}(x) = G(-8x + 3) - G(-4x + 3) \quad (5.13)$$

and  $b_{j,\ell} = \mathbf{b}\left(\frac{\ell}{2^j}\right)$ . This window function is used in Pietrobon *et al.* [121].

4. From the  $\mathcal{B}$ -spline function of order 3

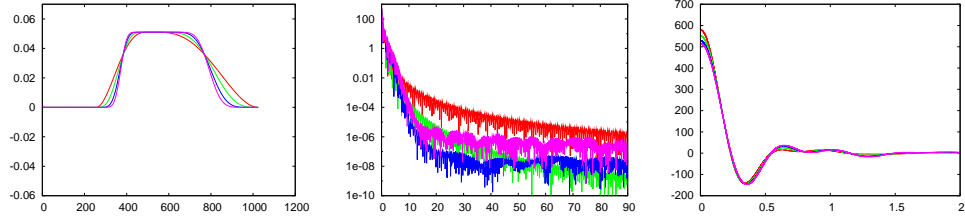
$$B_3(x) = \frac{1}{12}(|x - 2|^3 - 4|x - 1|^3 + 6|x|^3 - 4|x + 1|^3 + |x + 2|^3), \quad (5.14)$$

form  $\mathbf{b}(x) = \frac{3}{2}(B_3(2x) - B_3(x))$  and define  $b_{j,\ell} = \mathbf{b}\left(\frac{\ell}{2^j}\right)$ , which has support  $[0, 2^{j+1}]$ . This window function is used by Starck *et al.* [134].

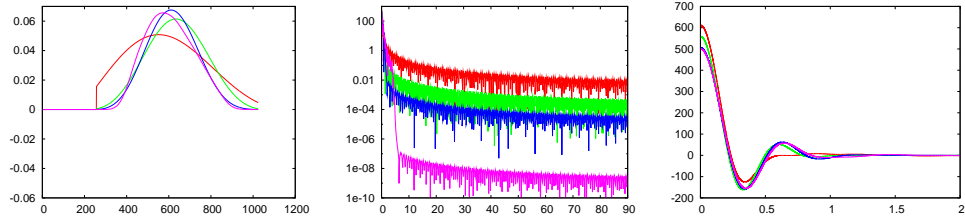
5. The Mexican hat wavelet on the sphere is the function the stereographic projection of which on the Euclidean plane is the usual Mexican hat wavelet. It has the following close expression depending on some positive scale parameter  $R$

$$\psi_R(\theta) \propto (1 - 2R^2 \tan^2(\theta/2)) \exp\{-2R^2 \tan^2(\theta/2)\}. \quad (5.15)$$

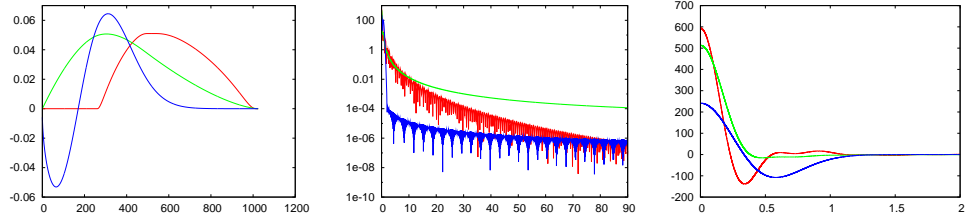
This wavelet is popular in the astrophysics community (see *e.g.* [83]). We have chosen  $R = 6.10^{-3}$  such that the spectral window is almost zero for  $\ell > 1024$ .



(a) Splines of order resp. 7, 15, 31 and 43.



(b) PSWFs localized in polar caps of 0.5, 1, 1.5 and 5 degree opening



(c) (red curve) Exponential function described in Eq. (5.13), (green curve)  $\mathcal{B}$ -spline function of Eq. (5.14) and (blue curve) Mexican hat described in Eq. (5.15).

F . 5.4 – In left column, the shape of the spectral windows as a function of  $\ell$ . In middle and right columns, the profile of the filters is plotted in the spatial domain as a function of  $\theta$  ( $\theta$  in degrees) with logarithmic and linear scales respectively, to illustrate both the decrease of the tail of the needlets far from the North pole and the shape of their first lobes. [fig. reproduite en couleur page 64]

### 5.2.2 $\mathbb{L}^2$ -concentration and variations

Our first attempt to achieve a good spatial localization of a needlet is to optimize a  $\mathbb{L}^2$ -norm based criterion, adapting to the sphere a problem that is well-known on the real line. In their seminal work in the 1960s and 1970s, Slepian and his collaborators studied the properties of prolate spheroidal wave functions (PSWFs) in the 1D case of real functions (see [42] and the references therein). PSWFs may be defined as functions with optimal energy concentration in the time domain, under some band-limitation constraint. Equivalently, they are the eigenfunctions of a time-frequency concentration kernel or the solutions of a Sturm-Liouville differential equation. The time-frequency concentration of PSWFs is understood in terms of continuous Fourier transform on  $\mathbb{R}$ . A discrete version of this theory, based on Fourier series coefficients, is derived in Slepian *et al.* [187].

In the last few years, Walter and coauthors exploited these 1D PSWFs to construct Slepian series (in [47]; see also [30]), and wavelets based on the best concentrated PSWF [48, 49].

#### Slepian functions on the sphere

On the sphere, we shall only consider the equivalent of Discrete PSWFs, which is studied in details by Simons, Dahlen and Wieczorek [41]. From a window function  $\{b_\ell\}$  with support  $I = [0, \ell^{max}]$ , define the axisymmetric function  $\psi$  by

$$\psi(\xi) = \sum_{\ell \in I} b_\ell L_\ell(\cos \theta). \quad (5.16)$$

The set of functions  $\psi$  of the form (5.16) is denoted  $B_I \subset \mathbb{L}^2(\mathbb{S}^2)$ . Given a spherical domain  $\Omega$ , consider the minimization, among non-zero functions in  $B_I$ , of the criterion

$$C_\Omega(\psi) = \frac{\int_{\mathbb{S}^2 \setminus \Omega} \psi^2(\xi) d\xi}{\int_{\mathbb{S}^2} \psi^2(\xi) d\xi} = 1 - \frac{\int_{\Omega} \psi^2(\xi) d\xi}{\int_{\mathbb{S}^2} \psi^2(\xi) d\xi}. \quad (5.17)$$

We call PSWF (by abuse of language) and denote  $\psi_\Omega^\star$  a normalized minimizer for  $C_\Omega(\psi)$ .

The criterion (5.17) has a simple expression when  $\Omega$  is axisymmetric. Consider the polar cap  $\Omega_{\theta_0} = \{\xi : \theta \leq \theta_0\}$  and define the coupling matrix  $\mathbf{D} = (D_{\ell, \ell'})_{\ell, \ell' \in I}$  by

$$D_{\ell, \ell'} = \frac{8\pi^2}{\sqrt{(2\ell+1)(2\ell'+1)}} \int_{\cos \theta_0}^1 L_\ell(z) L_{\ell'}(z) dz,$$

and the vector  $\bar{\mathbf{b}}(\psi) = (\bar{b}(\psi)_\ell)_{\ell \in I}$  by  $\bar{b}(\psi)_\ell = \sqrt{\frac{2\ell+1}{8\pi^2}} b_\ell$ . Then

$$C_\Omega(\psi) = 1 - \frac{\bar{\mathbf{b}}' \mathbf{D} \bar{\mathbf{b}}}{\|\bar{\mathbf{b}}\|^2} \quad (5.18)$$

and the minimization of  $C_\Omega$  becomes an eigenvalue problem. Note that the optimal solution depends on the opening  $\theta_0$ . As in the 1-dimensional case, the spectrum of  $\mathbf{D}$  exhibits a “step function” behaviour : denoting  $N = \text{tr } \mathbf{D}$  (the “Shannon number”), the matrix  $\mathbf{D}$  has about  $N$  eigenvalues very close to 1, and most of the others close to zero (see Fig.5.5, and Simons et al. [41] for details).

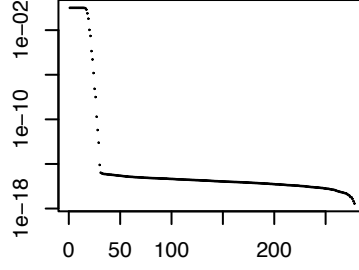


FIG. 5.5 – Eigenvalues of matrix  $\mathbf{D}$  with  $\theta_0 = 8^\circ$  and  $L = [129, 512]$ . In this case, Shannon number  $N \simeq 17$ .

### $\mathbb{L}^2$ -concentration criterion for window functions

We apply now the same idea to the design of window functions. We keep the same framework as above, except that  $I = [\ell^{\min}, \ell^{\max}]$  with  $\ell^{\max} \neq 0$ . In Figure 5.6 we plot the value of  $C_{\Omega_{\theta_0}}$  against  $\theta_0$  for  $\psi_{\Omega_{1^\circ}}^*$ ,  $\psi_{\Omega_{5^\circ}}^*$ ,  $\psi_{\Omega_{10^\circ}}^*$ . The lowest curve is the minimum of the criterion for all openings  $\theta_0$ . It is clear that there is no optimal function uniformly in  $\theta_0$ : the concentration criterion  $C_{\Omega_0}$  of each PSWF  $\psi_{\Omega_{\theta_1}}^*$  reaches the best possible value for  $\theta_0 = \theta_1$  only.

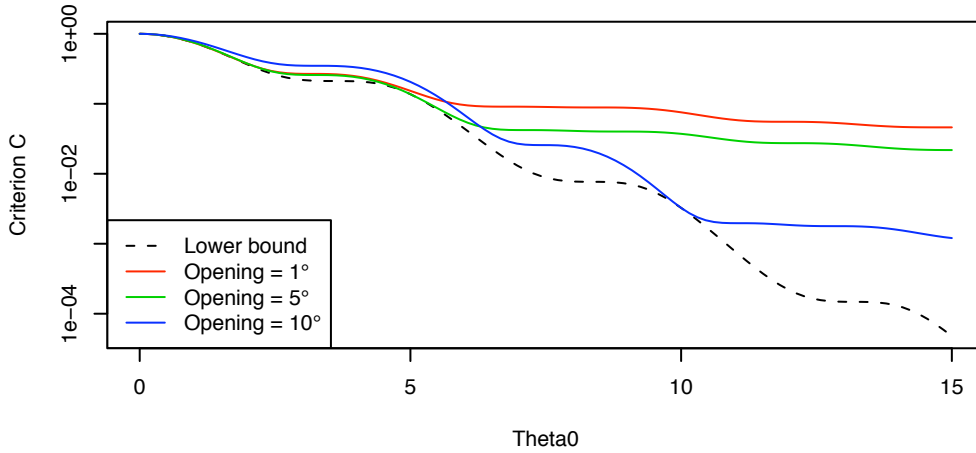


FIG. 5.6 – Localization for  $\mathbb{L}^2$ -energy criterion of PSWFs, band-limited into  $I = [33, 64]$ . The dashed line is the minimum of the criterion  $C_{\Omega_{\theta_0}}$  as a function of  $\theta_0$  and the other ones are the values of  $C_{\Omega_{\theta_0}}(\psi)$  evaluated at  $\psi = \psi_{\Omega_1}^*$ ,  $\psi_{\Omega_5}^*$  and  $\psi_{\Omega_{10}}^*$ . [fig. reproduite en couleur page 65]

When several eigenvalues of  $\mathbf{D}$  are extremely close to 1 (see Figure 5.5), it is computationally difficult to find the largest one and the associated eigenvector. In the case of  $\Omega$  a polar cap and  $\ell^{\min} = 0$ , one can advantageously solve the less degenerated eigenvalue problem associated with the Grünbaum differential equation [24] which has the same solutions as (5.17). We are not aware of an equivalent theory in the case  $\ell^{\min} > 0$ .

If  $\epsilon$  is a scalar of the order the machine precision, all vectors in  $V_\epsilon = \bigoplus_{\lambda \geq 1-\epsilon} \text{Ker}(\mathbf{D} - \lambda \mathbf{Id})$  have well localized spatial counterparts, but they are not necessarily positive (in spherical harmonic domain). Lack



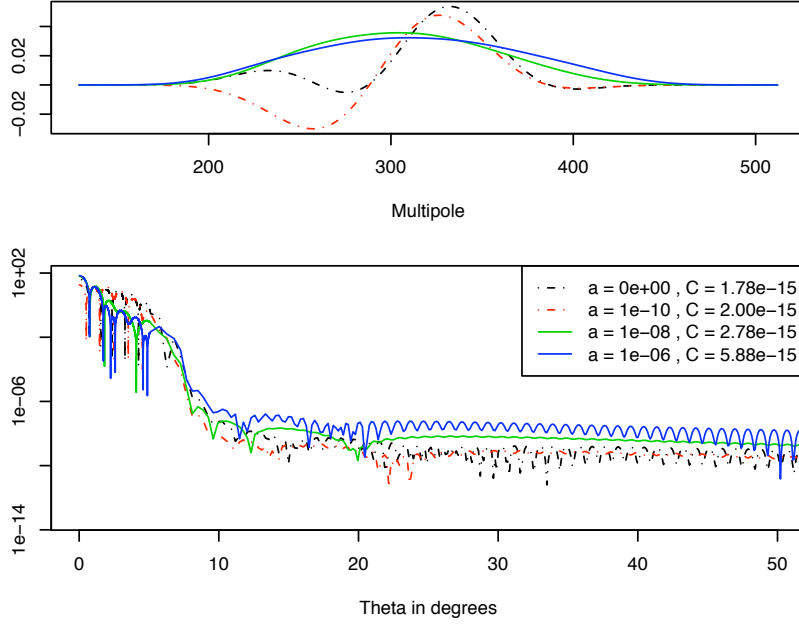


FIG. 5.7 – Effect of the smoothing on the spectral and spatial shapes of PSWFs associated to  $\theta_0 = 8$  degrees and  $L = [129, 512]$ .

of positivity is not acceptable for use in a needlet analysis-synthesis scheme where the spectral window is the square-root of the PSWF's window. To circumvent this difficulty, we introduce a modified coupling matrix  $\tilde{\mathbf{D}} = \mathbf{D} + a\mathbf{H}'\mathbf{H}$  where  $a > 0$  is a tuning parameter and  $\mathbf{H}$  is the tridiagonal second-order finite difference matrix. Window functions are now obtained as minimizers of  $\tilde{C}_\Omega(\psi) = 1 - \frac{\tilde{\mathbf{b}}'\tilde{\mathbf{D}}\tilde{\mathbf{b}}}{\|\tilde{\mathbf{b}}\|^2}$  instead of  $C_\Omega$ . The additional term favors non-oscillating functions among the vectors of  $V_\epsilon$  which are undistinguishable from their eigenvalues  $\lambda$ . Adding the “smoothing” term is expected not to alter the spatial localization of the filter. In practice, parameter  $a$  is selected to ensure ‘computational uniqueness’ of the smallest eigenvalue of  $\tilde{\mathbf{D}}$ . Solutions obtained by the numerical implementation of the minimization of  $\tilde{C}_{\Omega_{\theta_0}}$ ,  $\theta_0 = 8^\circ$  are displayed in Figure 5.7, with various values for the smoothing parameter  $a$ . Dashed lines correspond to the vector returned numerically as the “best” eigenvector of  $\mathbf{D}$  (associated to the greatest eigenvalue), and the best eigenvector of  $\tilde{\mathbf{D}}$  with parameter  $a$  chosen deliberately too small to ensure computationally uniqueness. Oscillating functions are indeed obtained. As  $a$  grows, the criterion selects non oscillating windows, two of which are plotted with solid lines. The loss measured by the increase of  $C_\Omega$  is displayed in the legend of the lower panel and appears extremely small. In our example, the energy outside  $\Omega$  for the needlet built from  $\tilde{C}$  takes the value  $2.78 \cdot 10^{-15}$ , whereas its minimal possible value is  $1.78 \cdot 10^{-15}$ .

### $\mathbb{L}^p$ -concentration criteria

A generalization of the Slepian concentration problem can be to consider other measures of concentration, such as  $\mathbb{L}^p$ ,  $p = 1, \dots, \infty$  instead of  $\mathbb{L}^2$ . The criterion defined in Eq. (5.17) becomes then

$$C_\Omega^{(p)}(\psi) = 1 - \frac{\|\psi \mathbf{I}_\Omega\|_p^p}{\|\psi\|_p^p} \quad (5.19)$$

where  $\|f\|_p^p = \int_{\mathbb{S}^2} |f(\xi)|^p d\xi$  if  $p \in [1, \infty)$  and  $\|f\|_\infty = \text{ess sup}_{\xi \in \mathbb{S}^2} |f(\xi)|$  for a spherical function  $f$ . Unlike

Slepian criterion  $C_\Omega = C_\Omega^{(2)}$ , these alternate criteria do not lead to simple eigenvalue problems. They could be numerically optimized but this is beyond the scope of this paper. However we compare in Section 5.3.1 this criterion to the original one  $C_\Omega$ .

### 5.2.3 Statistical criterion for optimal analysis with missing data

In this section, we consider the following framework : given an underlying random field  $X$  on  $\mathbb{S}^2$  to be analysed and two positive window functions  $W$  and  $D$  on  $\mathbb{S}^2$  multiplying the field, the aim is to get needlet coefficients of  $WX$  as close as possible to the coefficients computed from the uncorrupted field  $X$ .  $D$  specifies the relative importance of various regions on  $\mathbb{S}^2$  according, for instance, to the need for reliability in the needlet coefficients. For example,  $W$  can be a binary mask  $\mathbf{I}_\mathbb{W}$  setting the field to zero in some regions and  $D = \sum d_{\ell,m} Y_{\ell,m}$  can be  $\mathbf{I}_\mathbb{D}$  for a  $\mathbb{D} \subset \mathbb{S}^2$  of interest in which the analysis is to be done.

We shall assume statistical properties on the fields  $X, W, D$  and look for optimality of the filters on average.

**Assumption 5.1.** 1.  $X$  is a real-valued Gaussian zero mean isotropic square integrable random field on  $\mathbb{S}^2$ , with power spectrum  $C_\ell$ .  
2.  $W$  and  $D$  are deterministic elements of  $\mathbb{L}^2(\mathbb{S}^2)$ .

Implicitly,  $X$  is a measurable mapping from some  $(X, \mathbf{X}, \mathbf{P})$  into  $(\mathbb{L}^2(\mathbb{S}^2), \mathbf{H})$ ,  $\mathbf{H}$  being the Borel  $\sigma$ -field of  $\mathbb{L}^2(\mathbb{S}^2)$ . Let  $\mathbf{E}$  denote the expectation operator under  $\mathbf{P}$ . Recall that under Assumption 5.1, the covariance function on the field  $X$  is well defined and is given by

$$\mathbf{E}[X(\xi)X(\xi')] = (4\pi)^{-1} \sum_{\ell \in \mathbb{N}} C_\ell L_\ell(\xi \cdot \xi').$$

It follows that  $\mathbf{E}[X(\xi)^2] = (4\pi)^{-1} \sum_{\ell \in \mathbb{N}} (2\ell + 1)C_\ell$ . Moreover, the multipole moments  $(a_{\ell,m})$  of  $X$  are complex Gaussian random variables. They are centered, independent up to the relation  $a_{\ell,m} = a_{\ell,-m}^*$  and satisfy  $\mathbf{E}(|a_{\ell 0}|^2) = \frac{1}{2} \mathbf{E}(|a_{\ell,m}|^2) = C_\ell$ ,  $m \neq 0$ .

A first attempt in this direction is the derivation of an unbiased estimate of the spectrum from the multipole moments and the empirical power spectrum of the weighted sky  $XW$  defined by  $\hat{a}_{\ell,m} = \int_{\mathbb{S}^2} X(\xi)W(\xi)Y_{\ell,m}^*(\xi)d\xi$  and  $\hat{C}_\ell = \frac{1}{2\ell+1} \sum_m \hat{a}_{\ell,m}^2$  respectively. It is well-known (see [119, 89], see also the compact proof in Section 5.5) that

$$\mathbf{E}(\hat{C}_\ell) = \sum_{\ell' \in \mathbb{N}} \mathcal{M}_{\ell\ell'} C_{\ell'} \text{ with } \mathcal{M}_{\ell\ell'} = \sum_{0 \leq \ell'' \leq \ell + \ell'} \alpha_{\ell\ell'\ell''} \frac{2\ell'' + 1}{2\ell + 1} C_{\ell'}^W, \quad (5.20)$$

where the coefficients  $\alpha_{\ell\ell'\ell''}$  are defined by  $L_\ell L_{\ell'} = \sum_{0 \leq \ell'' \leq \ell + \ell'} \alpha_{\ell\ell'\ell''} L_{\ell''}$ . Note that the coupling matrix  $\mathcal{M}$  depends on  $W$  only through its ‘power spectrum’  $C_\ell^W$ . If  $\mathcal{M}$  is invertible, then  $(\mathcal{M}^{-1}(\hat{C}_{\ell'}))$  provides an unbiased estimate of  $(C_\ell)$ .

Let’s now derive a criterion to design a window function  $\mathbf{b} = (b_{\ell^{\min}}, \dots, b_{\ell^{\max}})$  which minimises the effect of missing data in a needlet analysis procedure. We focus on a single band smoothed field

(i.e. we fix one scale  $j$ ) and the dependence on  $j$  is implicit in the notations. For a collection of couple of indices, say  $(\ell_i, m_i)_{i=1, \dots, I}$ , we use  $\sum_{(\ell_i, m_i)_{i=1, \dots, I}}^*$  as a shorthand notation for the summation on  $\ell_i \in \mathbb{N}, m_i \in \{-\ell_i, \dots, \ell_i\}, i = 1, \dots, I$ .

Given  $\mathbf{b}$  and its associated smoothing operator  $\Psi = \sum_{\ell^{\min} \leq \ell \leq \ell^{\max}} b_\ell \Pi_\ell$ , the smoothed masked field is

$$\Psi X W(\xi) = \sum_{\ell \in I} b_\ell \int_{\mathbb{S}^2} X(\xi') W(\xi') L_\ell(\xi \cdot \xi') d\xi'.$$

Write  $\mathbf{E}[\Psi X(\xi)^2] = (4\pi)^{-1} \sum_\ell \sigma_\ell^2 b_\ell^2$  with  $\sigma_\ell^2 = (2\ell + 1)C_\ell$ . Let  $\epsilon$  denote the normalized difference field

$$\begin{aligned} \epsilon(\xi) &= \left( \mathbf{E}[\Psi X(\xi)^2] \right)^{-1/2} (\Psi X(\xi) - \Psi(XW)(\xi)) \\ &= \left( \sum_\ell \sigma_\ell^2 b_\ell^2 \right)^{-1/2} \sum_{(l,m)}^* b_\ell \bar{a}_{\ell,m} Y_{\ell,m}(\xi) \end{aligned} \quad (5.21)$$

where we have defined  $\bar{W} = 1 - W$ ,  $\bar{a}_{\ell,m} = \langle X\bar{W}, Y_{\ell,m} \rangle$ . Suppose that  $\{(\xi_k, \lambda_k)\}_{k \in K}$  provides an exact Gauss quadrature formula at a degree  $2\ell^{\max}$ . Define  $\beta_k$  and  $\beta'_k$  the needlet coefficients of  $X$  and  $XW$ , respectively and

$$\epsilon_k = \frac{\beta_k - \beta'_k}{\sqrt{\mathbf{E}(\beta_k^2)}}.$$

Those random variables are normalized errors on the needlet coefficients induced by the application of the weight function  $W$ . If both  $X$  and  $XW$  are in  $\mathbb{L}^2(\mathbb{S}^2)_{\ell^{\max}}$ , we easily check that  $\mathbf{E}(\beta_k^2) = \sqrt{\lambda_k} (4\pi)^{-1} \sum_\ell b_\ell^2 \sigma_\ell^2$ . Together with Proposition 5.3 it yields

$$\forall k \in K, \epsilon_k = \epsilon(\xi_k).$$

The dispersion of either the continuous field  $\epsilon(\xi)$  or the finite set  $\{\epsilon_k\}_{k \in K}$  is taken as a measure of quality for an analysis  $\Psi$ . This dispersion is not measured on the whole sphere, since the difference  $\epsilon$  must be important in the regions where  $W$  is far from 1, but weighted by the function  $D$ .

The coefficients  $\epsilon_k$  or their continuous version  $\epsilon$  are used in two ways. The first one introduces a “tolerance” threshold  $\alpha$  and counts the number of coefficients which are on average below this threshold. This measure of the efficiency of a filter in the presence of a mask is presented in Baldi et al. [151] and Pietrobon et al. [121] but its optimization was not considered. The second one considers the integrated square error of  $\epsilon$ , weighted by the function  $D$ . It leads to a quadratic criterion which is readily optimized.

The first criterion, writes, for a binary function  $D$ ,

$$E_{\mathbf{b}}(\alpha) = \frac{\sum_{k: D(\xi_k)=1} \mathbf{P}(|\epsilon_k| < \alpha)}{\#\{k : D(\xi_k) = 1\}}, \quad (5.22)$$

that is, the mean fraction of needlet coefficients corrupted by less than a normalized error  $\alpha \geq 0$ . For an arbitrary function  $D$ , a possible generalization of (5.22) is

$$E_{\mathbf{b}}(\alpha) = \frac{\sum_{k \in K} D(\xi_k) \mathbf{P}(|\epsilon_k| \leq \alpha)}{\sum_{k \in K} D(\xi_k)}.$$

In Subsection 5.3.2, we compare different windows using this criterion for a mask  $W$ .

Alternately, consider now the mean integrated square error (MISE)

$$R(\mathbf{b}) = \mathbf{E} \int_{\mathbb{S}^2} D(\xi) \|\epsilon(\xi)\|^2 d\xi \quad (5.23)$$

and define the optimal shape for the window  $\mathbf{b}$  as

$$\mathbf{b}^\star = \arg \min_{\|\mathbf{b}\|=1} R(\mathbf{b}). \quad (5.24)$$

Straightforward algebra leads to a close form expression of  $R(\mathbf{b})$  depending on  $\mathbf{b}$ , on the weight functions  $W$  and  $D$ , and on the power spectrum  $(C_\ell)_{\ell \in \mathbb{N}}$ . Let  $\bar{w}_{\ell,m}, d_{\ell,m}$  denote the multipole coefficients of the weight functions  $\bar{W}, D$ , respectively and define renormalized Wigner-3j coefficients

$$\begin{aligned} \begin{bmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{bmatrix} &:= \int_{\mathbb{S}^2} Y_{\ell,m}(\xi) Y_{\ell',m'}(\xi) Y_{\ell'',m''}^*(\xi) d\xi \\ &= (-1)^{m''} \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & -m'' \end{pmatrix} \end{aligned}$$

**Proposition 5.4.** *Under Assumption 5.1*

$$R(\mathbf{b}) = \frac{\mathbf{b}' \mathbf{Q} \mathbf{b}}{\mathbf{b}' \boldsymbol{\sigma} \mathbf{b}}$$

where  $\boldsymbol{\sigma} = \text{diag}((\sigma_\ell^2))$  and  $\mathbf{Q}$  is the matrix with entries

$$\begin{aligned} Q_{\ell\ell'} &= \sum_{m,m'} \sum_{(\ell_1,m_1)}^* C_{\ell_1} \\ &\quad \sum_{(\ell_i,m_i)_{i=2,3,4}}^* \bar{w}_{\ell_2 m_2} \bar{w}_{\ell_3 m_3}^* d_{\ell_4 m_4} \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m_1 & m_3 & m' \end{bmatrix}^* \begin{bmatrix} \ell & \ell_4 & \ell' \\ m & m_4 & m' \end{bmatrix}. \end{aligned} \quad (5.25)$$

If both  $W$  and  $D$  are axisymmetric the ten-tuple summations above reduce to a five-tuple one

$$\begin{aligned} Q_{\ell\ell'} &= \sum_m \sum_{\ell_1, \ell_2, \ell_3, \ell_4} C_{\ell_1} \bar{w}_{\ell_2 0} \bar{w}_{\ell_3 0}^* d_{\ell_4 0} \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m & 0 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m & 0 & m \end{bmatrix} \begin{bmatrix} \ell & \ell_4 & \ell' \\ m & 0 & m \end{bmatrix} \\ &= \sum_m A_{\ell\ell'm} D_{\ell\ell'm}. \end{aligned}$$

In the next section we shall give some illustrative examples of optimal spectral windows  $\mathbf{b}^\star$  in the particular axisymmetric case.

**Remark 5.3.** *As in the Slepian's problem, the design of an optimal filter reduces to an eigenvalue problem. In particular, if  $\sigma_\ell > 0$  for any  $\ell \in I$ , write  $b_\ell^\dagger = \sigma_\ell b_\ell$ . Let  $\mathbf{b}^{\dagger\star}$  be an eigenvector associated to the lowest eigenvalue of  $\mathbf{Q}^\dagger$ ,  $Q_{\ell\ell'}^\dagger = (\sigma_\ell \sigma_{\ell'})^{-1} Q_{\ell\ell'}$ . Then  $\mathbf{b}^\star := \boldsymbol{\sigma} \tilde{\mathbf{b}}^{\dagger\star} / \|\boldsymbol{\sigma} \tilde{\mathbf{b}}^{\dagger\star}\|$  is a solution of (5.24).*

**Remark 5.4.** *For those sums to be tractable, one has to assume that  $D, W, C_\ell$  have finite support in the frequency domain, i.e. that the windows  $D$  and  $W$  are smooth (or apodized) and  $C_\ell = 0$  for large enough  $\ell$ .*

**Remark 5.5.** *The matrix  $\mathbf{Q}$  being a second-order moment for the random field  $X$ , it can also be approximated by a moment estimator using Monte-Carlo experiments. This remark is of important practical interest as we are mostly concerned with non zonal masks.*

## 5.3 Examples, numerical results

### 5.3.1 Comparison of filters for various criteria

In Section 5.2, we considered several criteria measuring the localization properties of filters, and derived explicit or computational optimization for some of them. In Table 5.8, we compare the scores reached by the filters displayed in Figure 5.4. The columns indexed by  $\mathbb{L}^2$ - $\theta$  list the values  $C_{\Omega_\theta}(\psi)$  defined in Eq. (5.17). More generally, the columns indexed by  $\mathbb{L}^p$ - $\theta$  correspond to the values  $C_{\Omega_\theta}^p(\psi)$  defined in Eq. (5.19). A column lists the values of  $1 - E(\alpha)$  defined in Eq. (5.22), applied with the mask Kp0 of Fig. 5.9 and a tolerance parameter  $\alpha = 10\%$  (see next subsection for more details). Those quantities are estimated using Monte-Carlo simulations. A last column gives, by way of illustration only, the value of the “uncertainty product”  $|\mu(\psi)|^{-1} \times \sigma_\xi(\psi) \times \sigma_\ell(\psi)$ , where

$$\mu(\psi) := \int_{\mathbb{S}^2} \xi |f(\xi)|^2 d\xi \quad ; \quad \sigma_\xi(\psi) := \sqrt{1 - \|\mu(\psi)\|^2} \quad \text{and} \quad \sigma_\ell(\psi) := \sum_{\ell \geq 0} \ell(\ell+1) b_\ell^2. \quad (5.26)$$

Narcowich and Ward [34] proved that  $\Delta_\xi(\psi) \times \Delta_\ell(\psi) \geq 1$ .

	$\mathbb{L}^2$ -0.5°	$\mathbb{L}^2$ -1°	$\mathbb{L}^2$ -1.5°	$\mathbb{L}^2$ -5°	$\mathbb{L}^1$ -0.5°	$\mathbb{L}^1$ -1°	$\mathbb{L}^1$ -1.5°
Spline, order 3	2.2e-02	5.2e-03	7.4e-04	9.8e-07	4.2e-01	2.2e-01	1.0e-01
Spline, order 7	4.0e-02	1.3e-02	2.0e-03	4.8e-08	5.0e-01	2.9e-01	1.3e-01
Spline, order 15	6.1e-02	2.5e-02	4.9e-03	4.0e-07	5.9e-01	4.0e-01	2.2e-01
Spline, order 21	7.2e-02	3.1e-02	7.1e-03	7.7e-06	6.2e-01	4.5e-01	2.7e-01
Prolate, cap 0.5°	<b>1.2e-02</b>	6.0e-03	3.4e-03	9.5e-04	8.5e-01	8.2e-01	8.0e-01
Prolate, cap 1°	6.7e-02	<b>4.3e-05</b>	5.8e-06	1.7e-06	<b>3.8e-01</b>	1.3e-01	1.2e-01
Prolate, cap 1.5°	1.2e-01	1.5e-03	<b>3.4e-07</b>	1.2e-08	4.3e-01	<b>5.3e-02</b>	<b>1.0e-02</b>
Prolate, cap 5°	1.1e-01	6.7e-03	6.5e-04	<b>5.7e-14</b>	5.0e-01	1.8e-01	6.8e-02
Exponential	1.8e-02	3.2e-03	1.0e-03	1.0e-05	4.4e-01	2.6e-01	1.9e-01
B-Spline	<b>1.1e-02</b>	1.3e-03	3.9e-04	1.3e-05	4.8e-01	3.3e-01	2.7e-01
Mexican hat	6.4e-01	1.1e-02	8.5e-07	7.3e-12	7.9e-01	8.8e-02	<b>7.1e-04</b>
	$\mathbb{L}^1$ -5°	$\mathbb{L}^\infty$ -0.5°	$\mathbb{L}^\infty$ -1°	$\mathbb{L}^\infty$ -1.5°	$\mathbb{L}^\infty$ -5°	1-E(0.1)	$ \mu ^{-1} \sigma_\xi \sigma_\ell$
Spline, order 3	1.5e-02	5.0e-02	1.9e-02	5.1e-03	6.4e-05	26%	2.7
Spline, order 7	1.7e-03	6.0e-02	2.7e-02	7.1e-03	1.2e-05	33%	3.1
Spline, order 15	2.3e-03	6.9e-02	3.3e-02	9.8e-03	7.0e-05	41%	3.7
Spline, order 21	1.0e-02	7.3e-02	3.7e-02	1.1e-02	2.7e-04	46%	4.1
Prolate, cap 0.5°	7.2e-01	5.1e-02	1.0e-02	5.6e-03	1.0e-03	65%	9.8
Prolate, cap 1°	1.1e-01	1.1e-01	<b>2.0e-03</b>	2.0e-04	5.0e-05	<b>15%</b>	3.1
Prolate, cap 1.5°	8.8e-03	1.3e-01	1.7e-02	<b>1.4e-04</b>	4.5e-06	17%	3.6
Prolate, cap 5°	<b>1.1e-06</b>	1.2e-01	2.2e-02	5.9e-03	<b>2.6e-08</b>	24%	3.6
Exponential	4.8e-02	<b>4.4e-02</b>	1.4e-02	5.7e-03	1.9e-04	27%	<b>2.7</b>
B-Spline	1.5e-01	<b>3.1e-02</b>	6.8e-03	2.5e-03	1.5e-04	21%	<b>1.2</b>
Mexican hat	1.6e-04	4.7e-01	7.9e-02	4.5e-04	1.4e-07	49%	3.0

FIG. 5.8 – Comparison of the eleven filters of Fig. 5.4, the nine first of which are band-limited in  $I=[256,1024]$ .

The PSWFs perform the best not only for the  $\mathbb{L}^2$  criterion which they optimize, but also in most cases for the criteria where the  $\mathbb{L}^2$  norm is replaced by  $\mathbb{L}^p$  ones,  $p = 1$  and  $p = \infty$ , with the same opening angles  $\theta_0$ . Although the Kp0 mask has many small cut areas all over the sphere, most of the 11 filters presented here allow to retain more than 60% of the outside-mask coefficients  $\beta_k$  if a 10% error due to the presence of the mask is accepted. The performance w.r.t. this criterion goes up to 85% for the

PSWF optimally concentrated in a cap of  $1^\circ$ . However, the choice of arbitrary value of  $\alpha$  has a major impact on the ranking of the filters. This point is investigated in the next subsection.

### 5.3.2 Robustness of needlets coefficients

In this Subsection, we illustrate the performances of various window functions using the criterion (5.22). We have run  $N = 30$  Monte-Carlo experiments to estimate the numerator of  $E_{\mathbf{b}}(\alpha)$ . The random fields  $X$  are drawn using the  $(C_\ell)$ -spectrum of the best-fitting model for the CMB estimated by the WMAP team [86]. The mask  $W$  was chosen as Kp0, displayed in Figure 5.9, which masks the galactic plane and many point sources. The band is  $I = [256, 1024]$ .

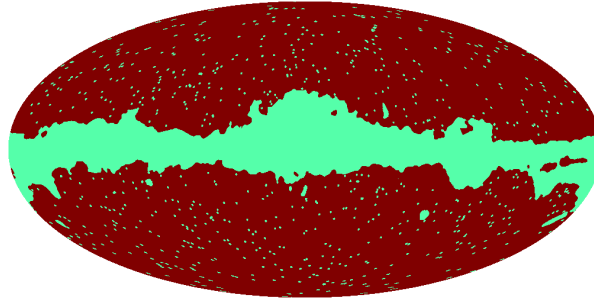


FIG. 5.9 – Kp0 mask.

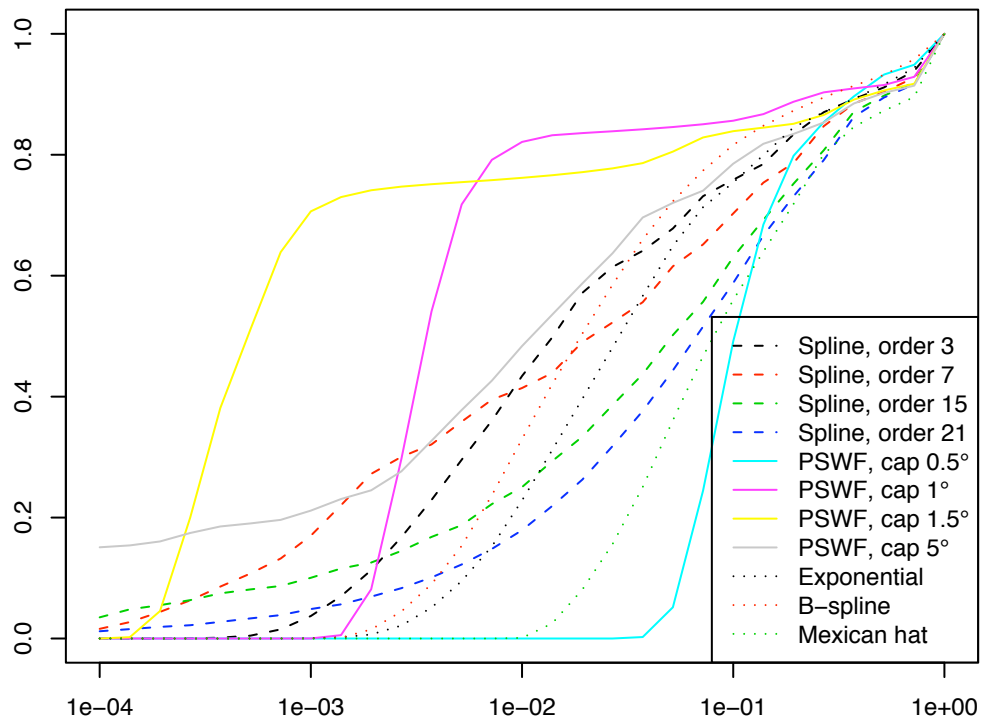
Figure 5.10 compares the increasing functions  $E_{\mathbf{b}}(\cdot)$  corresponding to various filters  $\mathbf{b}$ . There is no “uniformly best” (*i.e.* highest in the figure) needlet : some allow to retain more coefficients when the constraint imposed on the error is loose enough, but their efficiency decreases faster as  $\alpha$  goes to zero. Inspect *e.g.* the PSWF family.

### 5.3.3 Some MISE-optimal filters for axisymmetric weight functions

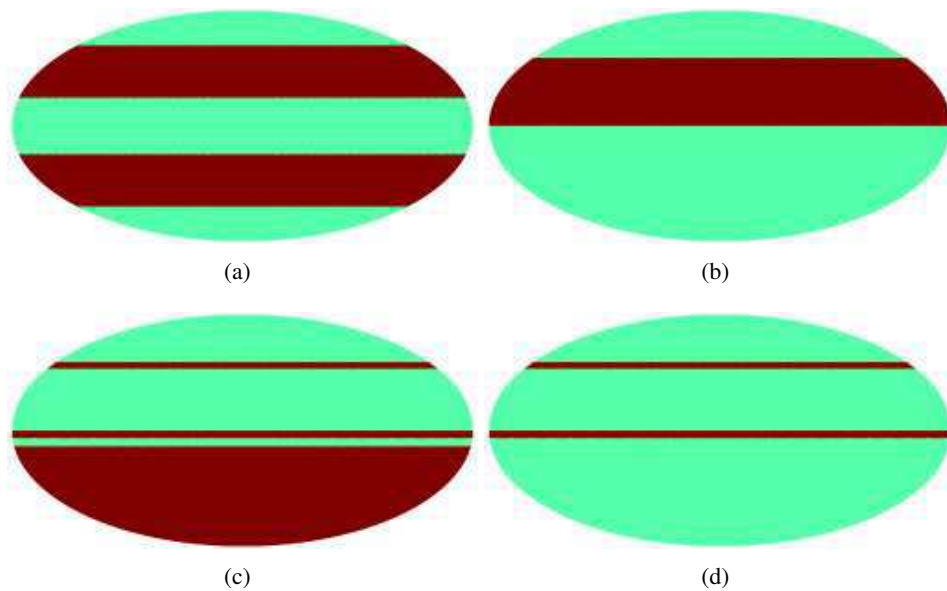
We present here the results of the optimization (5.24) in the case of axisymmetric weight functions  $W$ . For simplicity, the reconstruction weight function  $D$  is taken equal to  $W$ . We stick to the CMB spectrum of previous subsection. Figure 5.11 displays some of the masks  $W$  used in the experiments. This means that the data is available on the dark regions, and that its  $I = [\ell^{\min}, \ell^{\max}]$ -band-limited part has to be recovered in this area too. The apodization is simply a cosine-arch junction between 0 and 1, on a 2-degrees angular range.

On Figure 5.12 we have plotted the optimal filter in the  $R(\mathbf{b})$ -sense for the masks of Figure 5.11 together with different PSWFs. The criterion captures the symmetry of the mask (a) (the shape of the matrix  $\mathbf{Q}$  is a “checkerboard”), and the optimal filter is thus zero on all even (here) or all odd multipoles. The associated axisymmetric needlet  $\psi$  is symmetric w.r.t. the equatorial plane, and thus is well concentrated around both the North and the South poles. Such solutions are very sensitive to the modifications of the masks.

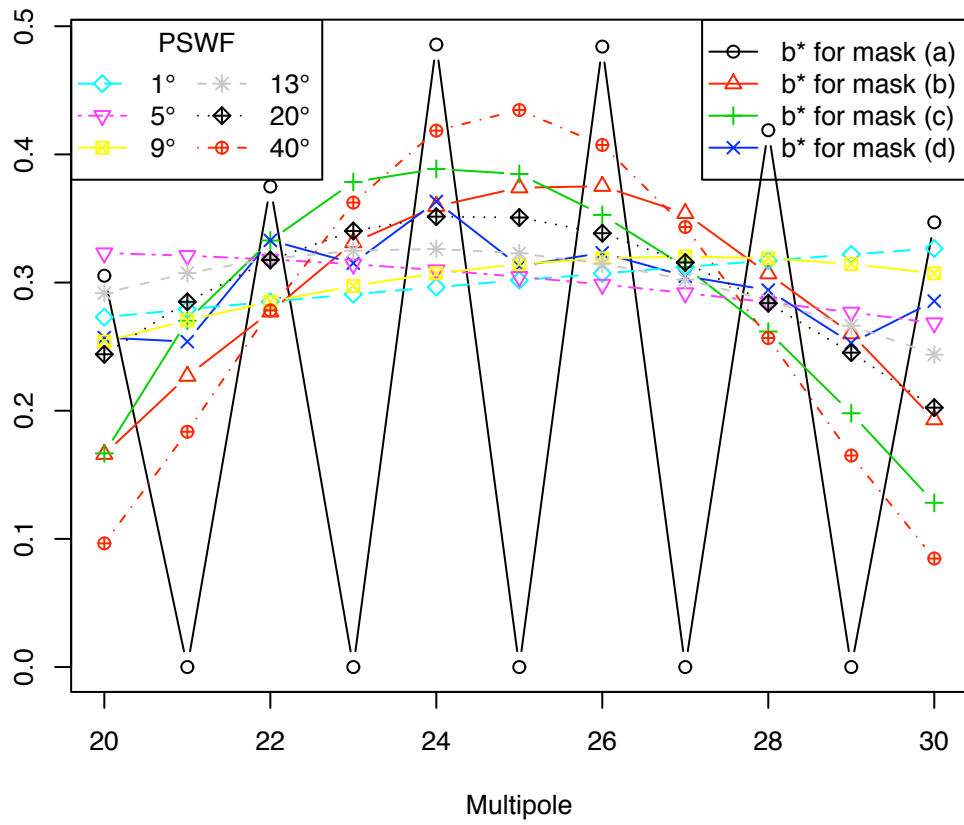
We conducted a small Monte-Carlo study to confirm the benefit of our approach. We have compared our best filters  $\mathbf{b}^*$  to PSWFs with different opening. On Figure 5.13, we show the box-plots of the distribution of the statistic  $R(\mathbf{h})$  for all those filters. The intercept of the horizontal line is  $R(\mathbf{b}^*)$  and



F . 5.10 – Proportion  $E_b(\alpha)$  of coefficients uncontaminated at tolerance level  $\alpha$ . [fig. reproduite en couleur page 65]

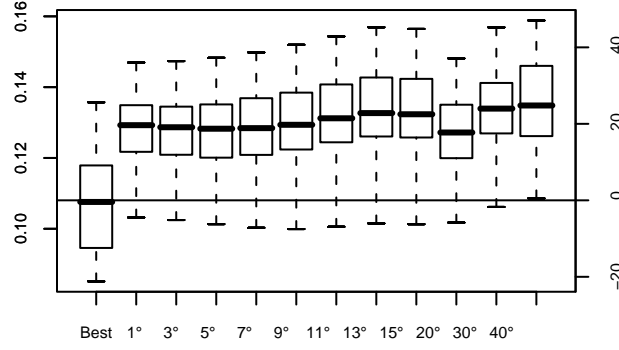


F . 5.11 – Four different apodized masks. The degree of apodization, measured as the width of the cosine-arch 0-1 junction is, 2 degrees.

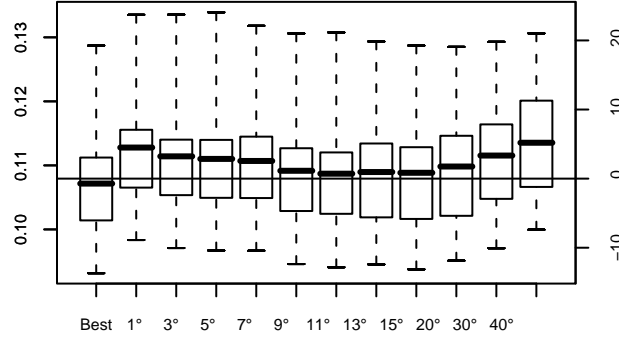


F . 5.12 – Shape of optimal window functions (plain lines) and PSWF (coloured and dashed lines) with various openings.





(a)



(b)

Fig. 5.13 – Empirical distribution of error  $R(\mathbf{b})$ . For Figure 5.13(a), the mask is Figure 5.11(a) and  $L = [5, 15]$ . For Figure 5.13(b), the mask is Figure 5.11(c) and  $L = [20, 30]$

the right vertical scale is for the relative error (in percent) with respect to this level.

Fig. 5.13(a) illustrates the strong benefit of a filter that captures the geometry of the mask. The relative improvement with respect to the best PSWF is of order 20%. It should be noted however that the shape of this optimal filter (described above) may lead to a misleading space-frequency picture. In some other cases, as shown in Figure 5.13(b), the relative improvement from the best PSWF to the best filter at all is very slight (a few percents). Here, the most favorable feature of our approach is that there is no tuning parameters (opening of the PSWF for instance, or the order of the splines window functions if they are taken as alternatives) to be found before the analysis.

## 5.4 Conclusion

A flexible way of analysing a field on the sphere in a space-frequency manner has been presented. It is based on the needlet construction of Narcowich *et al.* [33]. The analysis functions form a frame in the space a square-integrable functions on the sphere. Decompositions are essentially operating in the Spherical Harmonic domain, leading to fast implementations. Various criteria are used to design good spectral windows. This optimization can lead to decisive improvement in high precision applications

such as modern cosmology (CMB spectral estimation, component separation, etc.), where localized analysis is crucial.

**Acknowledgments.** We wish to thank Jacques Delabrouille for fruitful discussions motivating this work for CMB analysis. Numerical work was conducted using HEALPix [84].

## 5.5 Proofs

**Proof of Propositions 5.2** Firstly, using successively (5.2) and the quadrature assumption

$$\begin{aligned} \sum_{j \in \mathcal{J}, k \in K_j} |\beta_{j,k}|^2 &= \sum_{j \in \mathcal{J}} \sum_{k \in K_j} \lambda_{j,k} \left| \sum_{\ell=0}^{\ell_j^{\max}} b_{j,\ell} \Pi_\ell X(\xi_k) \right|^2 \\ &= \sum_{j \in \mathcal{J}} \sum_{\ell, \ell'=0}^{\ell_j^{\max}} b_{j,\ell} b_{j,\ell'} \int_{\mathbb{S}^2} \Pi_\ell X(\xi) \Pi_{\ell'} X(\xi) d\xi \\ &= \sum_{\ell \in \mathbb{N}} \sum_{j \in \mathcal{J}} (b_{j,\ell})^2 \|\Pi_\ell X\|^2. \end{aligned}$$

Using (5.9) and  $\|X\|^2 = \sum_{\ell} \|\Pi_\ell X\|^2$ , we get  $A\|X\|^2 \leq \sum_{j,k} |\beta_{j,k}|^2 \leq B\|X\|^2$ . Prove now that  $(\tilde{\psi}_{j,k})$  is the dual frame of  $(\psi_{j,k})$ . Write

$$\begin{aligned} \langle \tilde{\psi}_{j',k'}, \psi_{j,k} \rangle &= (\lambda_{j',k'} \lambda_{j,k})^{(1/2)} \sum_{\ell'=0}^{\ell_{j'}^{\max}} \sum_{\ell=0}^{\ell_j^{\max}} \tilde{b}_{j',\ell'} b_{j,\ell} \int_{\mathbb{S}^2} L_{\ell'}(\xi \cdot \xi_{j',k'}) L_\ell(\xi \cdot \xi_{j,k}) d\xi \\ &= (\lambda_{j',k'} \lambda_{j,k})^{(1/2)} \sum_{\ell=0}^{\ell_{j'}^{\max}} \tilde{b}_{j',\ell} b_{j,\ell} L_\ell(\xi_{j',k'} \cdot \xi_{j,k}) \end{aligned}$$

Then, for any  $j \in \mathbb{N}, k \in K_j$

$$\begin{aligned} \sum_{j',k'} \langle \tilde{\psi}_{j,k}, \psi_{j',k'} \rangle \psi_{j',k'} &= (\lambda_{j,k})^{1/2} \sum_{j' \in \mathcal{J}} \sum_{\ell=0}^{\ell_{j'}^{\max}} \tilde{b}_{j',\ell} b_{j,\ell} \sum_{\ell'=0}^{\ell_{j'}^{\max}} b_{j',\ell'} \int_{\mathbb{S}^2} L_\ell(\xi' \cdot \xi_{j,k}) L_{\ell'}(\xi' \cdot \xi) d\xi' \\ &= (\lambda_{j,k})^{1/2} \sum_{\ell=0}^{\infty} b_{j,\ell} \sum_{j' \in \mathcal{J}} \tilde{b}_{j',\ell} b_{j',\ell} L_\ell(\xi_{j,k} \cdot \xi) \\ &= \psi_{j,k}. \end{aligned}$$

The assertions (5.10) are a consequence of the dual frame property [see e.g. 9].

**Proof of Proposition 5.3** From Definition 5.2 of  $\psi_k$  and Eq (5.2)

$$\beta_k = \langle X, \psi_k \rangle = \sqrt{\lambda_k} \sum_{\ell} b_{\ell} \int_{\mathbb{S}^2} X(\xi) L_{\ell}(\xi, \xi_k) d\xi = \sqrt{\lambda_k} \Psi X(\xi_k).$$

**Proof of Eq. (5.20)**

$$\begin{aligned}
(2\ell + 1)\mathbf{E}(\hat{C}_\ell) &= \sum_{m=-\ell}^{\ell} \iint_{\mathbb{S}^2 \times \mathbb{S}^2} \mathbf{E} \{X(\xi)X(\xi')\} Y_{\ell,m}(\xi)Y_{\ell,m}(\xi')W(\xi)W(\xi')d\xi d\xi' \\
&= \iint_{\mathbb{S}^2 \times \mathbb{S}^2} \left\{ \sum_{\ell' \in \mathbb{N}} C_{\ell'} L_{\ell'}(\xi \cdot \xi') \right\} L_\ell(\xi \cdot \xi') W(\xi)W(\xi') d\xi d\xi' \\
&= \sum_{\ell' \in \mathbb{N}} C_{\ell'} \sum_{0 \leq \ell'' \leq \ell + \ell'} \alpha_{\ell\ell'\ell''} \iint_{\mathbb{S}^2 \times \mathbb{S}^2} L_{\ell''}(\xi \cdot \xi') W(\xi)W(\xi') d\xi d\xi' \\
&= \sum_{\ell' \in \mathbb{N}} C_{\ell'} \sum_{0 \leq \ell'' \leq \ell + \ell'} \alpha_{\ell\ell'\ell''} (2\ell'' + 1) C_{\ell''}^W
\end{aligned}$$

**Proof of Proposition 5.4** As  $X\bar{W} = \sum_{(\ell_1, m_1)}^* a_{\ell_1 m_1} Y_{\ell_1 m_1} \sum_{(\ell_2, m_2)}^* \bar{w}_{\ell_2 m_2} Y_{\ell_2 m_2}$ ,

$$\bar{a}_{\ell, m} := \langle X\bar{W}, Y_{\ell, m} \rangle = \sum_{(\ell_i, m_i)_{i=1,2}}^* a_{\ell_1 m_1} \bar{w}_{\ell_2 m_2} \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{bmatrix}.$$

Together with  $\mathbf{E}[a_{\ell, m} a_{\ell', m'}^*] = C_\ell \delta_{\ell\ell'} \delta_{mm'}$  it yields

$$\mathbf{E}[\bar{a}_{\ell, m} \bar{a}_{\ell', m'}^*] = \sum_{(\ell_i, m_i)_{i=1,2,3}}^* C_{\ell_1} \bar{w}_{\ell_2 m_2} \bar{w}_{\ell_3 m_3}^* \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m_1 & m_3 & m' \end{bmatrix}^* \quad (5.27)$$

Combining (5.21) and (5.27) we get

$$\begin{aligned}
R(\mathbf{b}) &= \left( \sum_{\ell \in \mathbb{N}} \sigma_\ell^2 h_\ell^2 \right)^{-1} \mathbf{E} \int_{\mathbb{S}^2} \sum_{(\ell_4, m_4)}^* d_{\ell_4 m_4} Y_{\ell_4 m_4}(\xi) \left| \sum_{(l, m)}^* h_\ell \bar{a}_{\ell, m} Y_{\ell, m}(\xi) \right|^2 d\xi \\
&= \left( \sum_{\ell \in \mathbb{N}} \sigma_\ell^2 h_\ell^2 \right)^{-1} \sum_{\ell, \ell' \in \mathbb{N}} h_\ell h_{\ell'} \sum_{mm'} \mathbf{E}[\bar{a}_{\ell, m} \bar{a}_{\ell', m'}^*] \sum_{(\ell_4, m_4)}^* d_{\ell_4 m_4} \int_{\mathbb{S}^2} Y_{\ell_4 m_4}(\xi) Y_{\ell, m}(\xi) Y_{\ell', m'}^*(\xi) d\xi \\
&= \left( \sum_{\ell \in \mathbb{N}} \sigma_\ell^2 h_\ell^2 \right)^{-1} \sum_{\ell, \ell' \in \mathbb{N}} h_\ell h_{\ell'} Q_{\ell\ell'}.
\end{aligned}$$

If  $W$  is axisymmetric,

$$\begin{aligned}
\mathbf{E}[\bar{a}_{\ell, m} \bar{a}_{\ell', m'}^*] &= \sum_{\ell_1, m_1}^* C_{\ell_1} \sum_{\ell_2, \ell_3} \bar{w}_{\ell_2 0} \bar{w}_{\ell_3 0} \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & 0 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m_1 & 0 & m' \end{bmatrix} \\
&= \delta_{m, m'} \sum_{\ell_1, \ell_2, \ell_3} C_{\ell_1} \bar{w}_{\ell_2 0} \bar{w}_{\ell_3 0} \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m & 0 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m & 0 & m \end{bmatrix} =: A_{\ell\ell' m} \quad (5.28)
\end{aligned}$$

where we used the fact that  $\begin{bmatrix} \ell & \ell' & \ell'' \\ m & 0 & m'' \end{bmatrix} = 0$  if  $m \neq m''$  and that  $w_{l_3 0}$  and  $\begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m & 0 & m \end{bmatrix}$  are real.

If  $D$  is axisymmetric and with  $D_{\ell\ell' m} := \sum_{\ell_4} d_{\ell_4 0} \begin{bmatrix} \ell & \ell_4 & \ell' \\ m & 0 & m \end{bmatrix}$

$$Q_{\ell\ell'} = \sum_m \sum_{(\ell_i, m_i)_{i=1,2,3}} C_{\ell_1} \bar{w}_{\ell_2 m_2} \bar{w}_{\ell_3 m_3}^* \begin{bmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m \end{bmatrix} \begin{bmatrix} \ell_1 & \ell_3 & \ell' \\ m_1 & m_2 & m \end{bmatrix}^* \sum_{\ell_4 \in \mathbb{N}} d_{\ell_4 0} \begin{bmatrix} \ell & \ell_4 & \ell' \\ m & 0 & m \end{bmatrix}.$$

## Chapitre 6

# CMB power spectrum estimation using wavelets

Ce chapitre reproduit l'article du même nom, G. Faÿ, F. Guillaoux, M. Betoule, J.-F. Cardoso, J. Delabrouille, M. Le Jeune, *Physical Review D* (2008)\*. Seules certaines notations ont été modifiées pour la cohérence du manuscrit.

**Abstract.** Observations of the Cosmic Microwave Background (CMB) provide increasingly accurate information about the structure of the Universe at the recombination epoch. Most of this information is encoded in the angular power spectrum of the CMB. The aim of this work is to propose a versatile and powerful method for spectral estimation on the sphere which can easily deal with nonstationary uncorrelated noise and multiple experiments with various specifications. In this paper, we use needlets (wavelets) on the sphere to construct natural and efficient spectral estimators for partially observed and beamed CMB with non stationary noise. In the case of a single experiment, we compare this method with Pseudo- $C_\ell$  methods. The performance of the needlet spectral estimators (NSE) compares very favorably to the best Pseudo- $C_\ell$  estimators, over the whole multipole range. On simulations with a simple model (CMB + uncorrelated noise with known variance per pixel + mask), they perform uniformly better. Their distinctive ability to aggregate many different experiments, to control the propagation of errors and to produce a single wide-band error bars is highlighted. The needlet spectral estimator is a powerful, tunable tool which is very well suited to angular power spectrum estimation of spherical data such as incomplete and noisy CMB maps.

### Sommaire

6.1	Framework . . . . .	106
6.1.1	Observation model . . . . .	106
6.1.2	Definition and implementation of a needlet analysis . . . . .	106
6.1.3	Distribution of the needlet coefficients . . . . .	107
6.1.4	Mask and beam effects . . . . .	108
6.2	The needlet spectral estimators (NSE) . . . . .	109
6.2.1	Smooth spectral estimates from a single map . . . . .	109
6.2.2	Smooth spectral estimates from multiple experiments . . . . .	110
6.2.3	Parameters of the method . . . . .	111

\*<http://dx.doi.org/10.1103/PhysRevD.78.083013>

6.3	Monte Carlo studies . . . . .	114
6.3.1	Single map with a mask and inhomogeneous noise . . . . .	115
6.3.2	Aggregation of multiple experiments . . . . .	119
6.4	Discussion . . . . .	125
6.5	Conclusion . . . . .	126
6.6	Appendices . . . . .	126
6.6.1	Pseudo- $C_\ell$ estimators . . . . .	126
6.6.2	What “noise spectrum” means . . . . .	127
6.6.3	Variance estimation by aggregation of experiments with independent heteroscedastic noise . . . . .	127

---

## Introduction

The estimation of the temperature and polarization angular power spectra of the Cosmic Microwave Background (CMB) is a key step for estimating the cosmological parameters. Cosmological information is encoded in the huge data sets (time order scanning data or high resolution maps) provided by ground-based, balloon-borne or satellite experiments.

In the ideal case of noiseless and full sky experiments, angular power spectrum estimation is a straightforward task. The empirical spectrum of the outcome of a Gaussian stationary field  $X$ , given by

$$\widehat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \langle X, Y_{\ell m} \rangle^2, \quad (6.1)$$

where  $(Y_{\ell m})$  denote the usual spherical harmonics, also is the *maximum likelihood* estimator of the power spectrum of  $X$ . It is efficient in the sense that its variance reaches the Cramér-Rao lower bound.

CMB maps are however more or less strongly contaminated by foregrounds and instrumental noises, depending on the wavelength, angular frequency  $\ell$  and the direction of observation. Ground-based experiments cover small parts of the sky while space missions (COBE, WMAP and, in the near future, Planck) provide full sky maps of the CMB, but still contaminated with galactic residuals. Then, the plain estimate (6.1) is no longer efficient nor even unbiased. To circumvent the non-stationarity of actual observations, the main ingredients for the spectral estimation used by, for instance, the WMAP collaboration [86] and also in most other analysis, are broadly the following ones. Usually, some part of the covered sky is blanked to remove the most emissive foregrounds or the most noisy measurements. This amounts to applying a *mask* or more generally a weight function to the sky. Most of the emissive foregrounds can be subtracted using some component separation procedure (see *e.g.* Leach *et al.* [100] for comparison methods with Planck-like simulated data). Even the best foreground-subtracting maps require masking a small fraction of the sky. Missing or masked data makes the optimal estimation of the power spectrum a much harder task. In particular, it breaks the diagonal structure of the covariance of the multipole moments  $a_{\ell,m} := \langle X, Y_{\ell,m} \rangle$  of any stationary component. The Maximum-likelihood estimation of the spectrum in the pixel domain has a numerical complexity that scales as  $N^3$  and requires the storage of a  $N^2$  matrices. This is untractable for high resolution experiments such as WMAP or Planck ( $N \approx 13 \cdot 10^6$ ). Nevertheless, for very low  $\ell$ 's ( $\ell \leq 30$ ), ML estimation in the pixel domain can be performed on downgraded resolution maps; see Tegmark [137] and Bond *et al.* [58], for instance. At higher  $\ell$ 's, a sub-optimal method based on the *Pseudo- $C_\ell$*  (PCL)

gives quite satisfactory results in terms of complexity and accuracy (Hivon *et al.* [89]). It debiases the empirical or (pseudo) spectrum from the noise contribution and deconvolves it from the average mask effect. It works in the spherical harmonic domain, uses fast spherical harmonic transforms (SHT) and scales as  $N^{3/2}$ . The available pixels can be weighted according to the signal to noise ratio (SNR) at any given point. For signal-dominated frequencies (low  $\ell$ 's), the data are uniformly weighted; it yields the Pseudo- $C_\ell$  estimator with uniform weights (PCLU). At noise-dominated frequencies (high  $\ell$ 's), each pixel is weighted by the inverse of the variance of the noise (PCLW estimator). The WMAP collaboration used uniform weights for  $\ell \leq 500$ , the inverse of the noise variance for  $\ell > 500$  (see [86, Section 7]) for its three-year release. Efstathiou [74] showed that the PCLW estimator is statistically equivalent to the ML estimator in the low SNR limit, which is usually the case at high  $\ell$ 's. In the same paper he proposed an hybrid method with a smooth transition between the two PCL regimes. Finally, when several maps are available, it is worth considering *cross-power spectra* between different channels since noise is usually uncorrelated from channel to channel (see [87, A1.1] or [122]).

Other estimation procedures do not fit in any of the two categories above. Among them, the spectral estimation from time ordered data by Wandelt *et al.* [144] or Gibbs sampling and Monte Carlo Markov chain methods such as MAGIC or Commander, see Eriksen *et al.* [77]. Those last methods try to estimate the complete posterior joint probability distribution of the power spectrum through sampling, which in turn can provide point estimates of the spectrum but also covariance estimates, etc. Recently, the multi-taper approach has been imported from the time series literature to the field of spherical data by Wieczorek, Simons and Dahlen [51, 182]. The goal of this approach is to provide an estimation of a localized power spectrum, under noiseless or stationary noisy assumption.

In this paper, we focus on spectral estimation of the global power spectrum, in a frequentist framework. We consider spectral estimation at small angular scales, *i.e.* in the range of multipoles where the cost of ML estimation is prohibitive. We compare our method to PCL methods. We adopt somehow realistic models that include partial coverage of the sky, symmetric beam convolution, inhomogeneous and uncorrelated additive pixel noise and multiple experiments.

Localized analysis functions such as wavelets are natural tools to tackle non-stationarity and missing data issues. There are different ways to define wavelets (in the broad sense of space-frequency objects) on the sphere, and our choice is to use the needlets, the statistical properties of which have received a recent rigorous treatment [149, 151] and which have already been applied successfully to cosmology [105, 121, 68].

Needlets benefit from perfect (and freely adjustable) localization in the spherical harmonic domain, which enables their use for spectral estimation. Moreover, the correlation between needlet coefficients centered on two fixed directions of the sky vanishes as the scale goes to infinity, *i.e.* as the needlet concentrate around those points. The spatial localization is excellent. This property leads to several convergence results and motivates procedures based on the approximation of decorrelation between coefficients. In this contribution, we define and study a new angular power spectrum estimator that uses the property of localization of the wavelets in both spatial and frequency domains.

In the case of a single experiment with partial coverage and inhomogeneous noise, the needlet-based estimator deals straightforwardly with the variation of noise level over the sky, taking advantage of their localization in the pixel domain. Moreover, it allows a joint spectral estimation from multiple experiments with different coverages, different beams and different noise levels. The proposed method mixes observations from all experiments with spatially varying weights to take into account the local noise levels. The resulting spectral estimator somehow mimics the maximum likelihood estimator

based on all the experiments.

The paper is organized as follows. In Section 6.1, we present the observation model and recall the basics of needlet analysis and the properties of the needlet coefficients which are the most relevant for spectral estimation. In Section 6.2, we define the needlet spectral estimators (NSE) in the single-map and multiple-maps frameworks. In Section 6.3, we present results of Monte Carlo experiments which demonstrate the effectiveness of our approach. In Section 6.4, we summarize the strong and weak points of our method and outline the remaining difficulties.

## 6.1 Framework

### 6.1.1 Observation model

Let  $T$  denote the temperature anisotropy of the CMB emission. For the sake of simplicity, we consider the following observation model

$$X(\xi_k) = W(\xi_k) ((B * T)(\xi_k) + \sigma(\xi_k)Z_k), \quad k = 1, \dots, N \quad (6.2)$$

where  $(\xi_k)$  is a collection of pixels on the sphere,  $W$  denotes a (0-1)-mask or any weight function  $0 \leq W \leq 1$ ,  $B$  denotes the instrumental beam. An additive instrumental noise is modelled by the term  $\sigma(\xi_k)Z_k$  with the assumption that  $(Z_k)$  is an independent standard Gaussian sequence. Further, we assume that  $\sigma$ ,  $W$  and  $B$  are known deterministic functions and that  $B$  is axisymmetric. Typically, the variance map  $\sigma^2$  writes  $\sigma^2(\xi_k) = \sigma_0^2/N_{\text{obs}}(\xi_k)$ , where  $N_{\text{obs}}$  is referred to as the *hit map*, that is,  $N_{\text{obs}}(\xi_k)$  is the number of times a pixel in direction  $\xi_k$  is seen by the instrument. We assume that the observations have been cleaned from foreground emissions or that those emissions are present but negligible outside the masked region. When observations from several experiments are jointly considered, the model becomes

$$X_e(\xi_k) = W_e(\xi_k) ((B_e * T)(\xi_k) + \sigma_e(\xi_k)Z_{k,e}) \quad k = 1, \dots, N, \quad e = 1, \dots, E \quad (6.3)$$

where  $e$  indexes the experiment. The CMB sky temperature  $T$  is the same for all experiments but the instrumental characteristics (beam, coverage) differ (see for example Table 6.2), and the respective noises can usually be considered as independent.

### 6.1.2 Definition and implementation of a needlet analysis

We recall here the construction and practical computation of the needlet coefficients. Details can be found in Guillaux *et al.* [25]; see also Narcowich *et al.* [33].

Needlets are based in a decomposition of the spectral domain in bands or ‘scales’ which are traditionally index by an integer  $j$ . Let  $b_{j,\ell}$  be a collection of window functions in the multipole domain, with maximal frequencies  $\ell_j^{\max}$  (see Figure 6.1 below). Consider some pixelization points  $\xi_{j,k}, k = 1, \dots, N_j$ , associated with positive weights  $\lambda_{j,k}, k = 1, \dots, N_j$  which enable exact discrete

integration (quadrature) for spherical harmonics up to degree  $2\ell_j^{max}$ , that is, equality

$$\int_{\mathbb{S}^2} Y_{\ell,m}(\xi) d\xi = \sum_{k=1}^{N_j} \lambda_{j,k} Y_{\ell,m}(\xi_{j,k})$$

holds for any  $\ell, m$  such that  $\ell \leq 2\ell_j^{max}$ ,  $|m| \leq \ell$ . Needlets are the axisymmetric functions defined by

$$\psi_{j,k}(\xi) = \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\ell_j^{max}} b_{j,\ell} L_\ell(\xi \cdot \xi_{j,k}), \quad (6.4)$$

where  $L_\ell$  denote the Legendre polynomial of order  $\ell$  normalized according to the condition  $L_\ell(1) = \frac{2\ell+1}{4\pi}$ . For proper choices of window functions  $\{b_{j,\ell}\}_j$ , the family  $\{\psi_{j,k}\}_{k,j}$  is a *frame* on the Hilbert space of square-integrable functions on the sphere  $\mathbb{S}^2$ . In a  $B$ -adic scheme, it is even a *tight frame* [33]. Though redundant, tight frames are complete sets which have many properties reminiscent of orthonormal bases (see *e.g.* [9], chap.3).

For any field  $X$  on the sphere, the coefficients  $\gamma_{j,k} := (\lambda_{j,k})^{-1/2} \langle X, \psi_{j,k} \rangle$  are easily computed in the spherical harmonic domain as made explicit by the following diagram

$$\begin{array}{ccc} \{X(\xi_k)\}_{k=1,\dots,N} & \xrightarrow{\text{SHT}} & a_{\ell,m} \\ & & \Downarrow \times \\ \{\gamma_{j,k}\}_{k=1,\dots,N_j} & \xleftarrow{\text{SHT}^{-1}} & b_{j,\ell} a_{\ell,m} \end{array} \quad (6.5)$$

Double arrows denotes as many operations (*e.g.* spherical transforms) as bands. The initial resolution must be fine enough to allow an exact computation of the Spherical Harmonics Transform up to degree  $\ell_j^{max}$ . If, say, the HEALPix pixelization is used, the `nside` parameter of the original map determines the highest available multipole moments and, in turn, the highest available band  $j$ .

### 6.1.3 Distribution of the needlet coefficients

A square-integrable random process  $X$  on the sphere is said to be centered and stationary (or isotropic) if  $\mathbf{E}(X(\xi)) = 0$ ,  $\mathbf{E}(X(\xi)^2) < \infty$  and  $\mathbf{E}(X(\xi)X(\xi')) = (4\pi)^{-1} \sum_{\ell} C_\ell L_\ell(\xi \cdot \xi')$ , with  $C_\ell$  referred to as the angular power spectrum of  $X$ . The next proposition summarizes the first and second order statistical properties of the needlet coefficients of such a process. They are the building blocks for any subsequent spectral analysis using needlets.

**Proposition 6.1.** *Suppose that  $X$  is a stationary and centered random field with power spectrum  $C_\ell$ . Then the needlet coefficients are centered random variables and, for any 4-tuple  $(j, j', k, k')$*

$$\mathbf{Cov}[\gamma_{j,k}, \gamma_{j',k'}] = \sum_{\ell \geq 0} b_{j,\ell} b_{j',\ell} C_\ell L_\ell(\cos \theta) \quad (6.6)$$

where  $\theta = \theta(j, k, j', k')$  is the angular distance between  $\xi_{j,k}$  and  $\xi_{j',k'}$ . In particular

$$\mathbf{Var}[\gamma_{j,k}] = C^{(j)}. \quad (6.7)$$

where

$$C^{(j)} := (4\pi)^{-1} \sum_{\ell \geq 0} (b_{j,\ell})^2 (2\ell + 1) C_\ell. \quad (6.8)$$



In other words, the variance of the coefficients  $\gamma_{j,k}$  is the power spectrum of  $X$  properly integrated over the  $j$ -th band.

**Remark 6.1.** *It also follows from (6.6) that if the bands  $j$  and  $j'$  are non-overlapping (this is the case for any non-consecutive filters of Figure 6.1), all the pairs of needlet coefficients  $\gamma_{j,k}$  and  $\gamma_{j',k'}$  are uncorrelated and then independent if the field is moreover Gaussian.*

Suppose now that

$$X(\xi_k) = \sigma(\xi_k)Z_k, \quad k = 1, \dots, N,$$

is a collection of independent random variables with zero mean and variance  $\sigma^2(\xi_k)$ , where  $\sigma$  is a band-limited function. This is a convenient and widely used model for residual instrumental noise (uncorrelated, but non-stationary). Needlet coefficients are centered and, moreover, if the quadrature weights are approximately uniform ( $\lambda_k \simeq 4\pi/N_{\text{pix}}$ , as is the case of HEALPix) and  $\sigma$  is sufficiently smooth, then

$$\text{Cov}[\gamma_{j,k}, \gamma_{j',k'}] \simeq \int_{\mathbb{S}^2} \sigma^2(\xi) \psi_{j,k}(\xi) \psi_{j',k'}(\xi) d\xi.$$

We denote

$$n_{j,k}(\sigma) := \left( \int_{\mathbb{S}^2} \sigma^2(\xi) |\psi_{j,k}(\xi)|^2 d\xi \right)^{1/2}. \quad (6.9)$$

the standard deviation of the needlet coefficient of scale  $j$  centered on  $\xi_k$ . When the noise is homogeneous ( $\sigma$  is constant), it reduces to  $\frac{\sigma^2}{N} \sum_{\ell \geq 0} (2\ell + 1) (b_{j,\ell})^2$ .

### 6.1.4 Mask and beam effects

As already noticed in the Introduction, missing or masked data makes the angular power spectrum estimation a non trivial task. Simple operations in Fourier space such as debeaming become tricky. Needlets are also affected by both the mask and the beam. The effect on needlets of beam and mask can be approximated as described below. These approximations, which lead to simple implementations, are validated in numerical simulations in relatively realistic conditions in Section 6.3.

**Mask.** Recalling that the needlets are spatially localized, the needlet coefficients are expected to be insensitive to the application of a mask on the data if they are computed far away from its edges. Numerical and theoretical studies of this property can be found in Baldi *et al.* [151] and Guillaux *et al.* [25]. In practice, we choose to quantify the effect of the mask on a single coefficient  $\gamma_{j,k}$  by the loss induced on the  $\mathbb{L}^2$ -norm of the needlet  $\psi_{j,k}$ , *i.e.* a purely geometrical criterion. More specifically, needlet coefficients at scale  $j$  are deemed reliable (at level  $t_j$ ) if they belong to the set

$$\mathcal{K}_j(t_j) := \left\{ k = 1, \dots, N_j : \frac{\|W\psi_{j,k}\|_2^2}{\|\psi_{j,k}\|_2^2} \geq t_j \right\} \quad (6.10)$$

Parameter  $t_j$  is typically set to 0.99 or 0.95 for all bands. Note that  $t \mapsto \mathcal{K}_j(t)$  is decreasing,  $\mathcal{K}_j(0) = \{1, \dots, N_j\}$  and  $\mathcal{K}_j(1) = \emptyset$ . In practice, this set is computed by thresholding the map obtained by the convolution of the mask with the axisymmetric kernel  $\xi \mapsto \left( \sum_{\ell} b_{j,\ell} L_{\ell}(\cos \theta) \right)^2$ . This operation is easy to implement in the multipole domain.

**Beam.** Consider now the effect of the instrumental beam. Its transfer function  $B_\ell$  is assumed smooth enough that it can be approximated in the band  $j$  by its mean value  $B_j$  in this band, defined by

$$(B_j)^2 := (4\pi)^{-1} \sum_{\ell \geq 0} (2\ell + 1) (b_{j,\ell})^2 B_\ell^2. \quad (6.11)$$

In the following, the beam effect for spectral estimation is taken into account in each band. Indeed, with Definition (6.11), with no noise, no mask and a smooth beam, Eq. (6.7) translates to

$$\mathbf{Var} \left[ \frac{\gamma_{j,k}}{B_j} \right] \simeq C^{(j)}, \quad k = 1, \dots, N_j. \quad (6.12)$$

In other words, thanks to the relative narrowness of the bands and to the smoothness of the beam and CMB spectrum, the attenuation induced by the beam can be approximated as acting uniformly in each band and not on individual multipoles. Numerically, with typical beam values from WMAP or ACBAR experiments (see Table 6.2), the relative difference (statistical bias) between the goal quantity  $C^{(j)}$  and the estimated one  $\mathbf{Var}(\gamma_{j,k})/(B_j)^2$  remains under 1% for bands below  $j = 27$  ( $\ell^{max} = 875$ ) for WMAP-W, and below  $j = 39$ ,  $\ell^{max} = 2000$  for ACBAR.

## 6.2 The needlet spectral estimators (NSE)

### 6.2.1 Smooth spectral estimates from a single map

For any sequence of weights  $w_{j,k}$  such that  $\sum_{k=1}^{N_j} w_{j,k} = 1$  and for a clean (contamination-free), complete (full-sky) and non-convolved (beam-free) observation of the CMB, the quantity

$$\widehat{C}^{(j)} := \sum_{k=1}^{N_j} w_{j,k} (\gamma_{j,k})^2$$

is an unbiased estimate of  $C^{(j)}$ , a direct consequence of Proposition 6.1.

**Remark 6.2.** For uniform weights, this estimator is nothing but the estimator  $\widehat{C}_\ell$  from Eq. (6.1) binned by the window function  $(b_{j,\ell})^2$ . Indeed, (see diagram (6.5))

$$\widehat{C}^{(j)} = (4\pi)^{-1} \sum_{\ell \geq 0} (b_{j,\ell})^2 \sum_{m=-\ell}^{\ell} a_{\ell m}^2 \quad (6.13)$$

$$= (4\pi)^{-1} \sum_{\ell \geq 0} (b_{j,\ell})^2 (2\ell + 1) \widehat{C}_\ell \quad (6.14)$$

This is the uniformly minimum variance unbiased estimator of  $C^{(j)}$ . The so-called cosmic variance is the Cramér-Rao lower bound for estimation of the parameter  $C_\ell$  in the full-sky, noise-free case. Its expression simply is  $2C_\ell^2/(2\ell + 1)$ . Its counterpart for the binned estimator  $C^{(j)}$  in this ideal context is

$$V_j^{cosmic} = 2(4\pi)^{-2} \sum_{\ell} (b_{j,\ell})^4 (2\ell + 1) C_\ell^2 \quad (6.15)$$

Consider now the observation model (6.2). Up to the approximations of Section 6.1.4, one finds that

$$\widehat{C}^{(j)} := \frac{1}{(B_j)^2} \sum_{k \in \mathcal{K}_j(t_j)} w_{j,k} \left\{ (\gamma_{j,k})^2 - (n_{j,k})^2 \right\}. \quad (6.16)$$

is an unbiased estimate of  $C^{(j)}$  as soon as

$$\sum_{k \in \mathcal{K}_j(t_j)} w_{j,k} = 1. \quad (6.17)$$

The weights can further be chosen to minimize the mean-square error  $\mathbf{E} \left( \widehat{C}^{(j)} - C^{(j)} \right)^2$ , under the constraint (6.17). It amounts to setting the weights according to the local signal-to-noise ratio, which is non constant for non stationary noise. This is a distinctive advantage of our method that it allows for such a weighting in a straightforward and natural manner. In the case of uncorrelated coefficients, this optimization problem is easy to solve (using Lagrange multipliers) and is equivalent to maximizing the likelihood under the approximation of independent coefficients (see Appendix 6.6.3 for details). It leads to the solution

$$w_{j,k}(\overline{C}) := \left( \overline{C} + (n_{j,k})^2 \right)^{-2} \left[ \sum_{k' \in \mathcal{K}_j(t_j)} \left( \overline{C} + (n_{k',j})^2 \right)^{-2} \right]^{-1} \quad (6.18)$$

with  $\overline{C} = C^{(j)}$ . This is the unknown quantity to be estimated but it can be replaced by some preliminary estimate (for example the spectral estimate of Hinshaw *et al.* [86]). One can also iterate the estimation procedure from any starting point. The robustness of this method with respect to the prior spectrum is demonstrated at Section 6.3.1 (see Figure 6.6).

Those weights are derived under the simplifying assumption of independence of needlet coefficients. They can be used in practice because needlet coefficients are only weakly dependent. Precisely, for two fixed points on a increasingly fine grid  $\xi_k, \xi_{k'}$  and well-chosen window functions, the needlets coefficients  $(\gamma_{j,k}, \gamma_{j,k'})$  are asymptotically independent as  $j \rightarrow \infty$  (see [151]). Note that this property is shared by well-known Mexican Hat wavelets, as proved in [175].

## 6.2.2 Smooth spectral estimates from multiple experiments

Consider now the observation model described by Eq. (6.3) with noise independent between experiments. Using the approximations of Section 6.1.4, Eq. (6.3) translates to

$$\frac{\gamma_{j,k,e}(X)}{B_{j,e}} = \gamma_{j,k}(T) + \frac{n_{j,k}(\sigma_e)}{B_{j,e}} Z_{k,e} \quad (6.19)$$

in the needlet domain, for indexes  $k \in \mathcal{K}_{j,e}(t_j)$ , where  $Z_{k,e}$  are standard Gaussian random variables which are correlated within the same experiment  $e$  but independent between experiments. As explained in the single experiment case of Section 6.2.1, the coefficients are only slightly correlated. This justifies the use, in the angular power spectrum estimator, of the weights derived by the maximization of the likelihood with independent variables. The correlation between coefficients does not introduce any bias here but only causes loss of efficiency. As in the single-experiment case, only the coefficients sufficiently far away from the mask of at least one experiment are kept. Defining

$$\mathcal{K}_j = \cup_e \mathcal{K}_{j,e}(t_j).$$

the aggregated estimator is implicitly defined (see Appendix 6.6.3) by

$$\widehat{C}^{ML,(j)} = \sum_{k \in \mathcal{K}_j} \tilde{w}_{j,k} \left( \widehat{C}^{ML,(j)} \right) \left\{ \left( \tilde{\gamma}_{j,k} \right)^2 - \left( \tilde{n}_{j,k} \right)^2 \right\} \quad (6.20)$$

with, for any  $k$  in  $\mathcal{K}_j$

$$\tilde{\gamma}_{j,k} = \sum_e \omega_{j,k,e} \frac{\gamma_{j,k,e}}{B_{j,e}} \quad (6.21)$$

$$\tilde{n}_{j,k} := \left[ \sum_e \left( \frac{B_{j,e}}{n_{j,k}(\sigma_e)} \right)^2 \mathbf{I}(k \in \mathcal{K}_{j,e}(t_j)) \right]^{-1/2} \quad (6.22)$$

$$\omega_{j,k,e} := \left( \frac{B_{j,e}}{n_{j,k}(\sigma_e)} \right)^2 \mathbf{I}(k \in \mathcal{K}_{j,e}(t_j)) \left( \tilde{n}_{j,k} \right)^2 \quad (6.23)$$

and similarly to (6.18),

$$\tilde{w}_{j,k}(C) := \left( C + \left( \tilde{n}_{j,k} \right)^2 \right)^{-2} \left[ \sum_{k' \in \mathcal{K}_j} \left( C + \left( \tilde{n}_{j,k'} \right)^2 \right)^{-2} \right]^{-1} \quad (6.24)$$

Note that  $\sum_e \omega_{j,k,e} = 1$  and  $\sum_k \tilde{w}_{j,k} = 1$ . An explicit estimator is obtained by plugging some previous, possibly rough, estimate  $\bar{C}_j$  of  $C^{(j)}$  in place of  $C$  of Eq. (6.24). Eventually, the aggregated angular power spectrum estimator is taken as

$$\widehat{C}^{(j)} = \sum_{k \in \mathcal{K}_j} \tilde{w}_{j,k} \left( \bar{C}_j \right) \left\{ \left( \tilde{\gamma}_{j,k} \right)^2 - \left( \tilde{n}_{j,k} \right)^2 \right\}. \quad (6.25)$$

This expression can be interpreted in the following way. For any pixel  $k$  in  $\mathcal{K}_j$ , that is for any pixel where the needlet coefficient is reasonably uncontaminated by the mask for at least one experiment, compute an aggregated needlet coefficient  $\tilde{\gamma}_{j,k}$  by the convex combination (6.21) of the debeamed needlet coefficients from all available experiments. Weights of the combination are computed according to the relative local signal to noise ratio (including the beam attenuation). Finally, a spectral estimation is performed on the single map of aggregated coefficients, in the same way as in Section 6.2.1. Those coefficients are squared and translated by  $\tilde{n}_{j,k}$  to provide an unbiased estimate of  $C^{(j)}$ . Then all the available squared and debiased coefficients are linearly combined according to their relative reliability  $\tilde{w}_{j,k}(C)$  which is proportional to  $(C^{(j)} + (\tilde{n}_{j,k})^2)^{-1}$ . Figures 6.13 and 6.14 display the values of those weights (maps)  $\tilde{w}_{j,k}$  and  $\omega_{j,k,e}$  for a particular mixing of experiments. See Section 6.3.2 for details.

### 6.2.3 Parameters of the method

In this section, we discuss various issues raised by the choice of the parameters of the NSE method. Those parameters are : the shape of the spectral window function  $b_{j,\ell}$  in each band (or equivalently the shape of the needlet itself in the spatial domain), the bands themselves (*i.e.* the spectral support of each needlet) and the values of the thresholds  $t_j$  that define the regions of the sky where needlets coefficients are trusted in each band ; see Eq. (6.10). See Section 6.3.1 for a numerical investigation.

Band ( $j$ )	1	2	3	4	5	...	20	21	22	23	24	25
$\ell_j^{\min}$	2	11	21	31	41	...	401	451	501	551	601	651
$\ell_j^{\max}$	20	30	40	50	60	...	500	550	600	650	700	750
$\text{nside}_j$	16	16	32	32	32	...	256	512	512	512	512	512
$\theta_{0j}$	69	50	41	36	32	...	10.7	10.2	9.7	9.3	8.9	8.6

Band ( $j$ )	26	27	28	29	...	35	36	37	38	39
$\ell_j^{\min}$	701	751	801	876	...	1326	1426	1501	1626	1751
$\ell_j^{\max}$	800	875	950	1025	...	1475	1625	1750	1875	2000
$\text{nside}_j$	512	512	512	1024	...	1024	1024	1024	1024	1024
$\theta_{0j}$	8.3	8.0	7.7	7.4	...	6.1	5.8	5.6	5.4	5.2

TABLE 6.1 – Spectral bands used for the needlet decomposition in this analysis. Depending on  $\ell_j^{\max}$ , the needlet coefficient maps are computed using the HEALPix package, at different values of  $\text{nside}$ , given in the fourth line. The number of needlet coefficients in band  $j$  is then  $12(\text{nside}_j)^2$ . It is a kind of *decimated* implementation of the needlet transform. The last line gives the opening  $\theta_j$  (in degrees) chosen in Eq. 6.26 to define the PSWF from the *Prolate 2* family (see Section 6.3.1 for details).

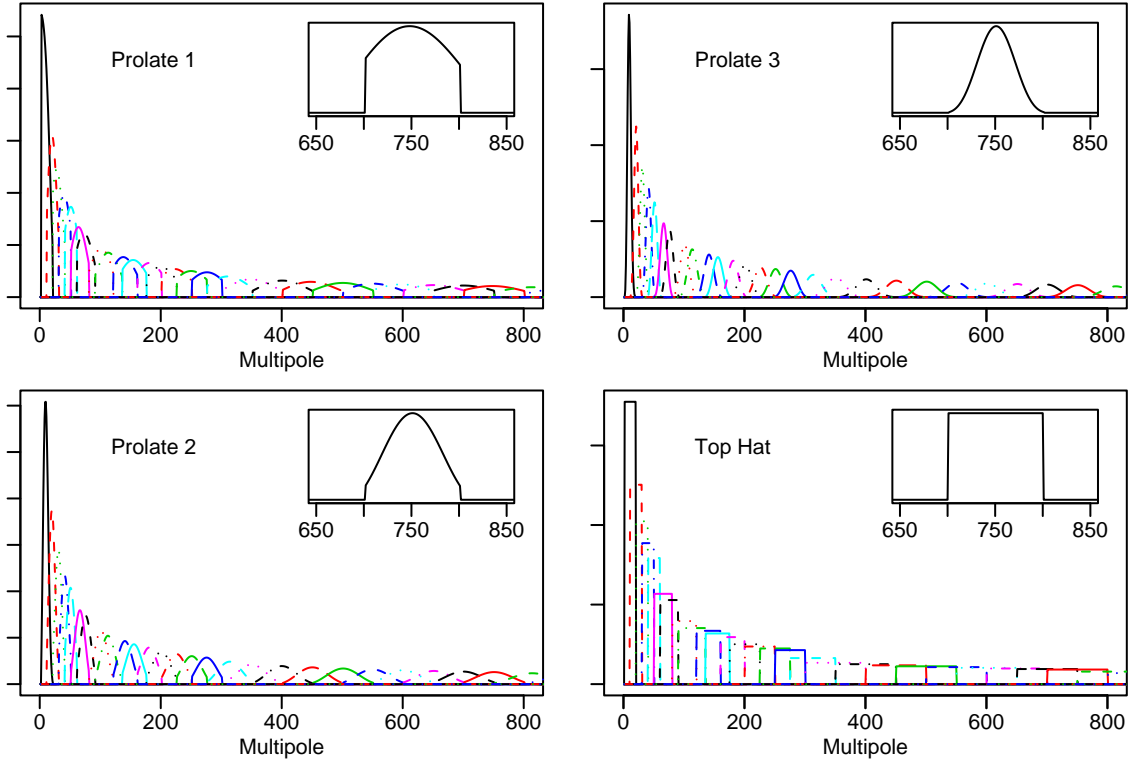
### Width and shape of the window functions

For spectral estimation, it is advisable to consider spectral window functions with relatively narrow spectral support, in order to reduce bias in the spectral estimation. The span of the summation in (6.4) can be fixed to some interval  $[\ell_j^{\min}, \ell_j^{\max}]$ . For our illustrations, the interval bands have been chosen to cover the range of available multipoles with more bands around the expected positions of the peaks of the CMB. The bands are described in Table 6.1.

It is well known however that perfect spectral and spatial localization cannot be achieved simultaneously (call it the uncertainty principle). In order to reduce the effect of the mask, we have to check that the analysis kernels are well localized. This leads to the optimization of some localization criteria. If we retain the best  $\mathbb{L}^2$  concentration in a polar cap  $\Omega_{\theta_j} = \{\xi : \theta \leq \theta_j\}$ , namely

$$(b_{j,\ell})_{\ell=\ell_j^{\min}, \dots, \ell_j^{\max}} = \arg \max_{\mathbf{b}} \frac{\int_{\Omega_{\theta_j}} \left| \sum_{\ell=\ell_j^{\min}}^{\ell_j^{\max}} \mathbf{b}_{\ell} L_{\ell}(\xi) \right|^2 d\xi}{\int_{\mathbb{S}^2} \left| \sum_{\ell=\ell_j^{\min}}^{\ell_j^{\max}} \mathbf{b}_{\ell} L_{\ell}(\xi) \right|^2 d\xi} \quad (6.26)$$

we obtain the analogous of prolate spheroidal wave function (PSWF) thoroughly studied in *e.g.* Slepian & Pollak [188] for the PSWF in  $\mathbb{R}$  and [41, 25] for PSWF on the sphere. In our simulations, we use PSWF needlets since they are well localized and easy to compute. Other criteria and needlets can be investigated and optimized, at least numerically; see Guilloux *et al.* [25] for details. The choice of the optimal window function in a given band is a non trivial problem which involves the spectrum itself, the characteristics of the noise and the geometry of the mask. Even if we restrict to PSWF as we do here, it is not clear how to choose the optimal opening  $\theta_j$  for each band  $j$ . We can use several rules of thumb based on approximate scaling relation between roughly  $B$ -adic bands and openings  $\theta_j$  that preserve some Heisenberg product or Shannon number. Figure 6.1 represents three families of PSWF needlets that are numerically compared below. Their spatial concentration is illustrated by Figure 6.2.



F . 6.1 – Four families of window functions that are used for the NSE and compared numerically in Section 6.3.1. There are three families of prolate spheroidal wave functions and one family of top-hat functions. All the families are defined on the same bands. Inset graphs show the window function in the 26th band. Each window function is normalized by the relation  $(4\pi)^{-1} \sum (b_{j,\ell})^2 (2\ell + 1) = 1$ . Then, if the angular power spectrum is flat,  $C_\ell \equiv C_0$ , then  $C^{(j)} \equiv C_0$  for all bands, according to (6.8).

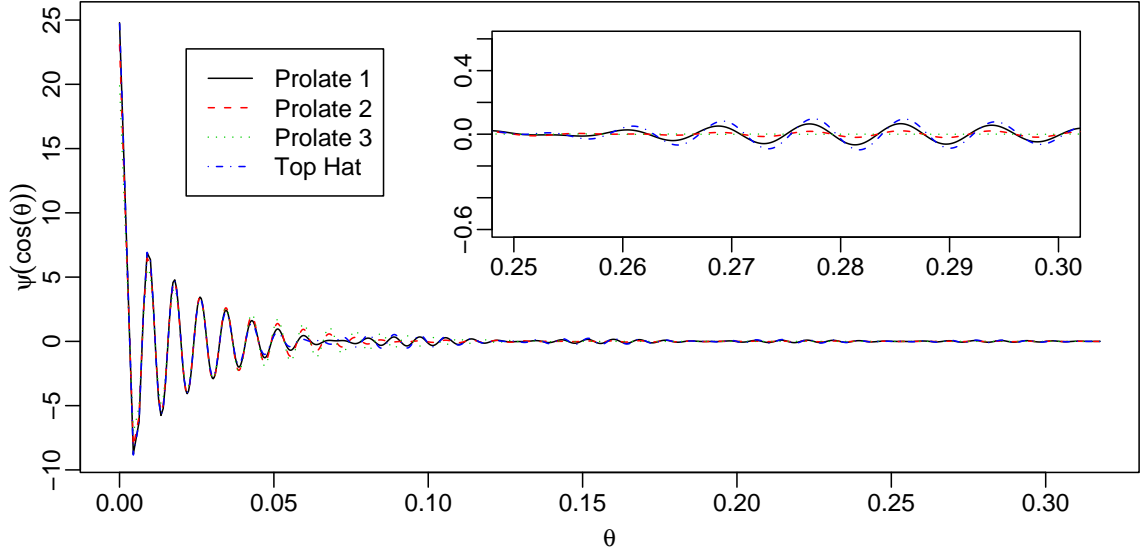


FIG. 6.2 – Angular profile of the four needlets associated to the window functions at the 26th band ( $701 \leq \ell \leq 800$ ) from the four families of Figure 6.1. We have plotted the axisymmetric profile  $\sum_{\ell} b_{\ell} L_{\ell}(\cos \theta)$  as a function of  $\theta$ . Needlets with “smoother” associated window profile (such as Prolate 3) need more room to get well localized, but are less bouncing than needlets with abrupt window function (such as top hat or Prolate 1)

### The choice of the needlets coefficients (mask)

Practically we want to keep as much information (*i.e.* as many needlet coefficients) as possible, and to minimize the effect of the mask. Using all the needlet coefficients regardless of the mask would lead to a biased estimate of the spectrum. It is still true if we keep all the coefficients outside but still close to the mask, keeping in mind that the needlets are not perfectly localized. On the other hand, getting rid of unreliable coefficients reduces the bias, but increases the variance. This classical trade-off is taken by choosing the threshold level  $t_j$  in the Definition (6.10) of the excluding zones. For multiple experiments, a different selection rule can be applied to each experiment, according to the geometry of the mask and the characteristics of the beam and the noise.

## 6.3 Monte Carlo studies

Recall that NSE spectral estimators are designed based on three approximations :

- one can neglect the impact of the mask on the needlet coefficients which are centered far enough from its edges ;
- one can neglect the variations of the beam and the CMB power spectrum over each band.
- the weights, which are optimal under the simplifying assumption of independent needlet coefficients, still provide good estimates for the truly weakly correlated needlet coefficients.

We carry out Monte Carlo studies to investigate, first the actual performance of the method on realistic data, and second the sensitivity of the method with respect to its parameters. Stochastic convergence

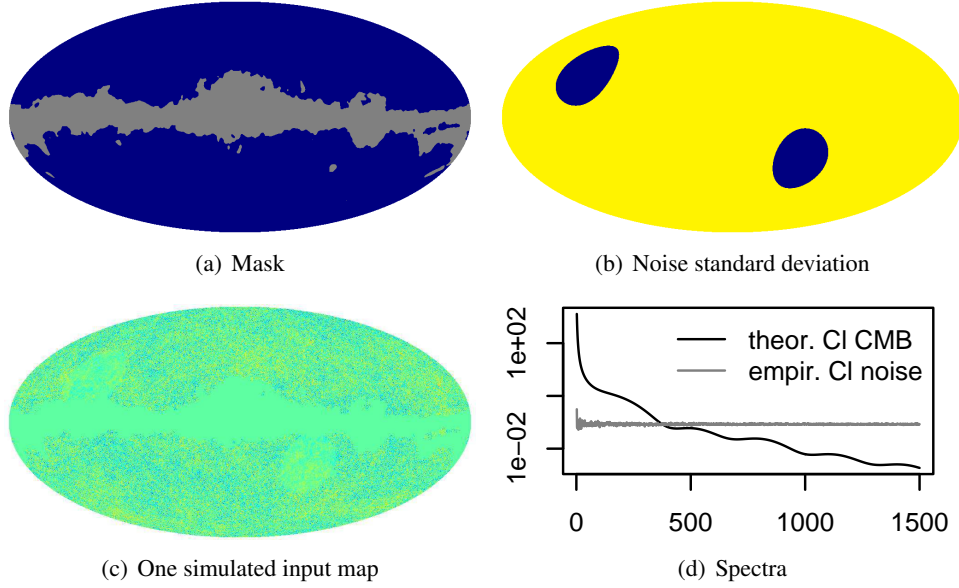


FIG. 6.3 – Simplified model of partially covered sky and inhomogeneous additive noise. This model is used to compare numerically the NSE estimator with PCL estimators and to assess the robustness or the sensitivity of the method with respect to its parameters. The mask is kp0. In CMB  $\mu K$  units, the standard deviation of the uncorrelated pixel noise is 75 in the two small circular patches and 300 elsewhere.

results under appropriate conditions is established in a companion paper [157].

### 6.3.1 Single map with a mask and inhomogeneous noise

In this section, we first consider model (6.2). According to Eqs (6.12) and (6.16), any beam can be taken into account easily in the procedure. Without loss of generality, we suppose here that there is no beam (or  $B$  is the Dirac function). The case of different beams is addressed in Section 6.3.2, see Table 6.2.

The key elements for this numerical experiment are illustrated by Fig. 6.3. We simulate CMB from the spectrum  $C_\ell$  given by the  $\Lambda$ CDM model that best fits the WMAP data. We use a Kp0 cut [53] for the mask and we take a simple non homogeneous noise standard deviation map (the SNR per pixel is 1.5 in two small circular patches and 0.4 elsewhere).

Using a mean-square error criterion, we first study the dependence of NSE performance on its free parameters. Then we compare NSE with methods based on Spherical Harmonic coefficients, known as pseudo- $C_\ell$  estimation and followed in Hinshaw *et al.* [86]. For the reader's convenience, the PCL procedure is summarized in Appendix 6.6.1.



## Mean-square error

We shall measure the quality of any estimator  $\widehat{C}^{(j)}$  of  $C^{(j)}$  by its mean-square error

$$\text{MSE}(\widehat{C}^{(j)}) = \mathbf{E} \left( \widehat{C}^{(j)} - C^{(j)} \right)^2 .$$

This expectation is estimated using 800 Monte Carlo replications. Roughly speaking, the MSE decomposes as an average estimation error and a *sampling variance*. The estimation error term is intrinsic to the method. Ideally, it should be used to compare the relative efficiency of concurrent approaches. The sampling variance term is the so-called *cosmic variance*. It is given by the characteristic of the spectrum and coming from the fact that we only have one CMB sky, and thus  $2\ell + 1$   $a_{\ell,m}$ 's to estimate one  $C_\ell$ . It is increased by the negative influence of the noise and the mask. This gives an error term intrinsic to the whole experiment. When the sky is partially observed (let  $f_{\text{sky}}$  denote the fraction of available sky) and for high  $\ell$ 's (or  $j$ 's), the cosmic variance must be divided by a factor  $f_{\text{sky}}$  leading to the following approximate Cramér-Rao lower bound at high frequencies

$$V_j^{\text{sample}} = f_{\text{sky}}^{-1} V_j^{\text{cosmic}} . \quad (6.27)$$

Including an homogeneous additive uncorrelated pixel noise with variance  $\sigma^2$ , the sample variance writes

$$2f_{\text{sky}}^{-1} \sum_{\ell=0}^{\infty} (b_{j,\ell})^4 (2\ell + 1) \left( C_\ell + \frac{4\pi}{N} \sigma^2 \right)^2$$

In a non-homogeneous context, no close expression for the sampling variance is available : Eq. (6.27) will serve as one reference. When comparing different window functions in the same band, it must be kept in mind that different estimators do not estimate the same  $C^{(j)}$  so that the sampling variances are not the same. In this case, we use the following normalized MSE

$$\frac{\text{MSE}(\widehat{C}^{(j)})}{f_{\text{sky}}^{-1} V_j^{\text{cosmic}}} \quad (6.28)$$

## Robustness with respect to parameter choice

This section looks into the robustness of NSE with respect to its free parameters.

First and as expected, the spectral estimation is very sensitive to the choice of the window functions. Even if we restrict to the PSWF, one has the freedom to choose a concentration radius  $\theta_j$  for each band. We compare the mean-square error of the estimation for various choices of  $\theta_j$  that lead to three of the window function families displayed in Figure 6.1. The second prolate family is obtained using the “rule of thumb” relation  $\theta_j = 2((\ell_j^{\min} + \ell_j^{\max})/2)^{-1/2}$ . The values of those opening angles are in Table 6.1. The first and third sequences of opening angles are the same with a multiplicative factor of 0.5 and 2, respectively. For the sake of comparison we also consider the top-hat window functions. Figure 6.4 shows the normalized MSE for those four “families” of needlets as a function of the band index. Notice the poor behavior of a non-optimized window function and the far better performance of the second prolate family in comparison with top-hat and Prolate 1 windows. Thus, in the following, we use this particular needlet family to study the sensitivity of the method with respect to the other parameters, and to compare NSE and PCL estimators.

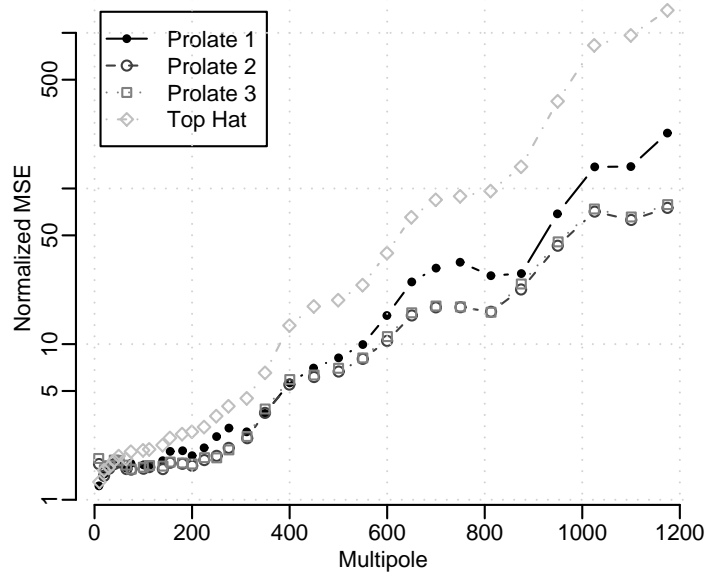


FIG. 6.4 – Comparison of the normalized MSE (6.28) of the needlet spectral estimators for the four families of spectral window functions displayed in Figure 6.1. The smoothness of the window function make the MSE smaller at high multipoles. At low multipoles, taking a too smooth function makes the needlet less localized and there is a loss of variance due to the smaller number  $\mathcal{K}_j$  of needlet coefficients that are combined.

Next, for the second family of PSWF, we compare the influence of the threshold value  $t_j \equiv t$  for  $t = 0.9, 0.95$  or  $0.98$ . Figure 6.5 shows that this choice within reasonable values is not decisive in the results of the estimation procedure. For very low frequencies ( $\ell \leq 100$ ), the conservative choice  $t = 0.98$  increases the variance since many needlets are contaminated by the mask and discarded. Qualitatively, in such variance dominated regimes, taking more coefficients (*e.g.*  $t = 0.9$ ) is adequate. However, we do not advocate the use of the NSE at low  $\ell$ 's where exact maximum likelihood estimation is doable. At higher  $\ell$ 's, there is roughly no difference between the  $t = 0.95$  and  $t = 0.98$  thresholds.

Finally, we check the robustness of the method against an imprecise initial spectrum. We take  $0.9C^{(j)}$  and  $1.1C^{(j)}$  as initial values  $\bar{C}^{(j)}$  and compare the results with the best possible initial value which is  $C^{(j)}$  itself. The relative difference between the results, displayed in Figure 6.6, does not exceed 1%.

### Pseudo- $C_\ell$ versus needlet spectral estimator

We compare the NSE estimator given by Eq. (6.16) with estimation based on the spherical Harmonic coefficients of the uniformly weighted map and of the  $1/\sigma^2$ -weighted map. The result is displayed in Fig. 6.7.

As expected, at low multipoles, where the SNR is higher, the uniform pseudo- $C_\ell$  estimator performs better than the weighted pseudo- $C_\ell$  estimator and, conversely, at high multipoles where the SNR is lower. Efstathiou [74] proved that the equal-weights pseudo- $C_\ell$  estimator is asymptotically Fisher-efficient when  $\ell$  goes to infinity. The behaviour of the needlet estimator is excellent : its performance

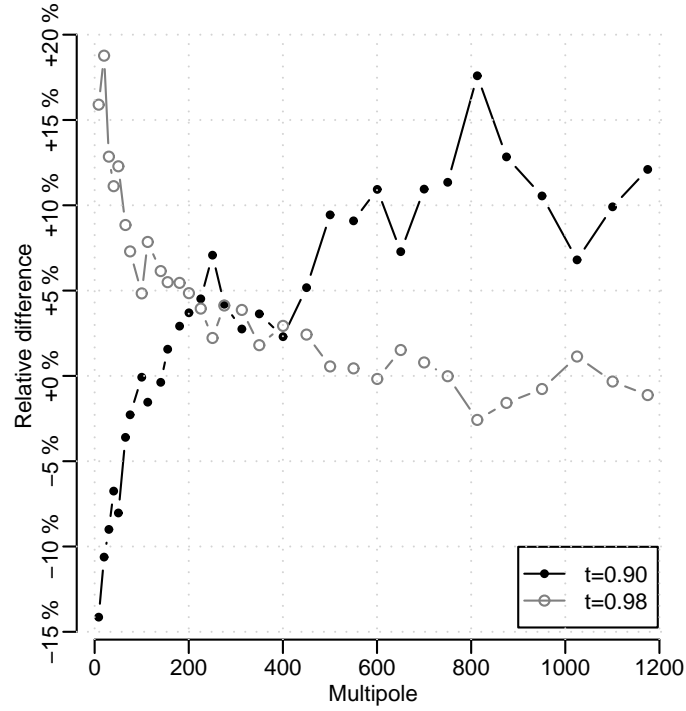


Figure 6.5 – Relative difference between the normalized MSE (6.28) of the needlet spectral estimation using thresholds 0.9 and 0.98, and the same with threshold 0.95. It highlights the fact that the estimation is not very sensitive to the value of this parameter, except at low  $\ell$ 's, where we do not advocate the use of the NSE. The window function family is “Prolate 2” from Figure 6.1.

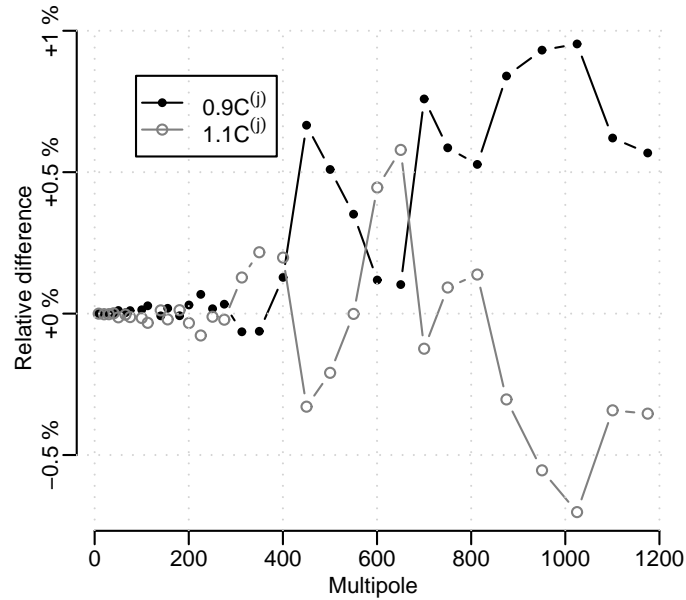


Figure 6.6 – Robustness of the NSE with respect to the initial value  $\bar{C}^{(j)}$  given to the weights formula (6.18). We have performed the whole estimation with  $\bar{C}^{(j)} = 0.9C^{(j)}$  and  $\bar{C}^{(j)} = 1.1C^{(j)}$ . This plot shows the relative difference between the normalized MSE under those initial values and the normalized MSE under the optimal initial value  $\bar{C}^{(j)} = C^{(j)}$ .

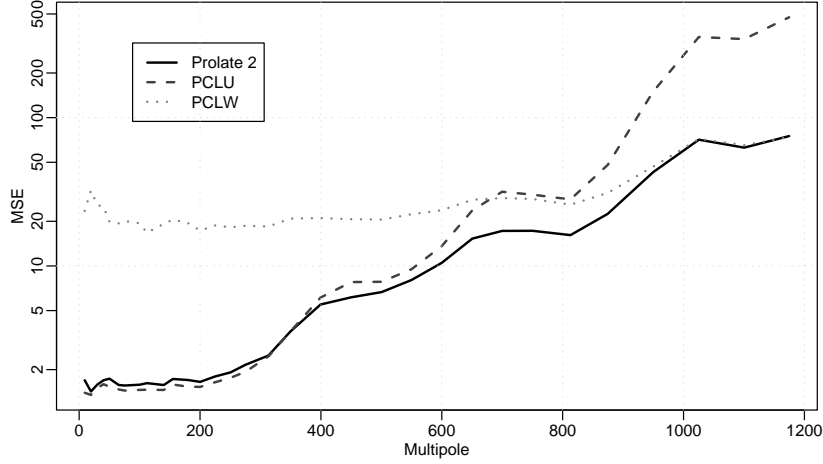


FIG. 6.7 – Comparison of the relative MSE (6.28) of the two PCL estimators (PCLU for flat weights, PCLW for inverse variance weights) with the NSE, for Prolate family 2. For  $300 \leq \ell \leq 1200$ , the NSE is uniformly better than the best of the two PCL methods. It should be noted that the NSE may be improved again by optimizing the window profiles and the thresholds  $t_j$  (e.g. by taking a lower threshold for low bands to reduce the variance, see Figure 6.5).

is comparable to the best of the two previous methods both at low and high multipoles. Thus, there is no need to choose arbitrary boundaries between frequencies for switching from one weighting to the other. The NSE estimator automatically implements a smooth transition between the two regimes and it does so quasi-optimally according to noise and mask characteristics. At low  $\ell$ 's one should optimize the window function (the characteristic angle of opening of the prolates) to broaden the range of optimality of the NSE.

Providing a  $C_\ell$  estimate with error bars is often not sufficient. Estimating the covariance matrix of the whole vector of spectral estimates is necessary for full error propagation towards, say, estimates of cosmological parameters. Figure 6.8 shows the values of the correlation matrix between the spectral estimates. In the idealistic case of a full sky noiseless experiment, the theoretical correlation matrix is tridiagonal because window functions we have chosen only overlap with their left and right nearest neighbours. The mask induces a spectral leakage, which is however reduced for the smoothest window function. This leakage is however compensated for by the selection of coefficients in  $\mathcal{K}_j(t_j)$  (see Eq (6.10)).

### 6.3.2 Aggregation of multiple experiments

Historically in CMB anisotropy observations, no single instrument provides the best measurement everywhere on the sky, and for all possible scales. In the early 90's, the largest scales have been observed first by COBE-DMR, complemented by many ground-based and balloon-borne measurements at higher  $\ell$ . Similarly, ten years later, WMAP full sky observations on large and intermediate scales have been complemented by small scale, local observations of the sky as those of Boomerang, Maxima, ACBAR or VSA.

The joint exploitation of such observations has been so far very basic. The best power spectrum is

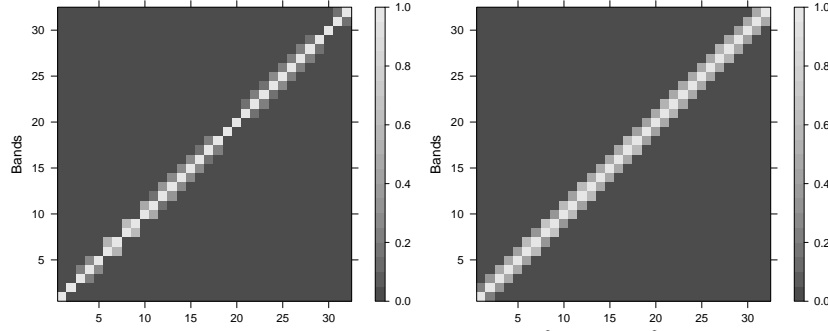


FIG. 6.8 – Absolute value of the correlation matrix of vector  $(\widehat{C}^{(1)}, \dots, \widehat{C}^{(32)})$ , which entries are defined by Eq. (6.16). It has been estimated in the context of Fig. 6.3 using two different families of window functions and 800 Monte Carlo replicates. This shows the difference between a family of PSWF (left panel) and a family of top-hat windows (right panel).

Experiment	Beam	nside	Noise level	$f_{\text{sky}}$
WMAP Q	31'	512	Given by the hit map	78.57 %
WMAP V	21'			
WMAP W	13'			
BOOM S	10'	1024	17.5 $\mu\text{K}$	2.80 %
BOOM D			5.2 $\mu\text{K}$	0.65 %
ACBAR	5'	2048 <sup>1</sup>	14.5 $\mu\text{K}$	1.62 %

TABLE 6.2 – Main parameters of the experiments to be aggregated. The beams are given in minutes of arc, nside refers to HEALPix resolution of the simulated maps, noise level is either a map computed from a hitmap and an overall noise level, or a uniform noise level per pixel (in  $\mu\text{K}$  CMB). Numbers quoted here are indicative of the typical characteristics of observations as those of WMAP, BOOMERanG and ACBAR, and are used for illustrative purposes only.

obtained by choosing, for each scale, the best measurement available, and discarding the others. One could, alternatively, average the measurements in some way, but the handling of errors is complicated in cases where a fraction of the sky is observed in common by more than one experiment.

Clearly, the data is best used if some method is devised that allows combining such complementary observations in an optimal way. In this section we present the results of a Monte Carlo study to illustrate the benefits of our method of aggregated spectral estimator.

We simulate observations following the model (6.3), with  $E = 6$  observed maps : 3 Kp0-masked maps with beams and noise-level maps according to WMAP experiment in bands Q, V and W respectively ; 3 maps with uniform noise, observed in patches the size of which are equivalent to BOOMERanG-Shallow, BOOMERanG-Deep and ACBAR observations respectively, and noise levels representative of the sensitivities of those experiments. Table 6.2 gives the key features of these experiments. Further details can be found in Bennett *et al.* [53] and Hinshaw *et al.* [88] for WMAP, Masi *et al.* [106] for BOOMERanG, [125] and Reichardt *et al.* [124] for ACBAR. However, we do not intend to produce fully-realistic simulations. Basically, no foregrounds are included in simulations (neither diffuse nor point sources) ; for ACBAR only the 3 sky fields of year 2002 are used ; and for WMAP only one detector is used for each band.

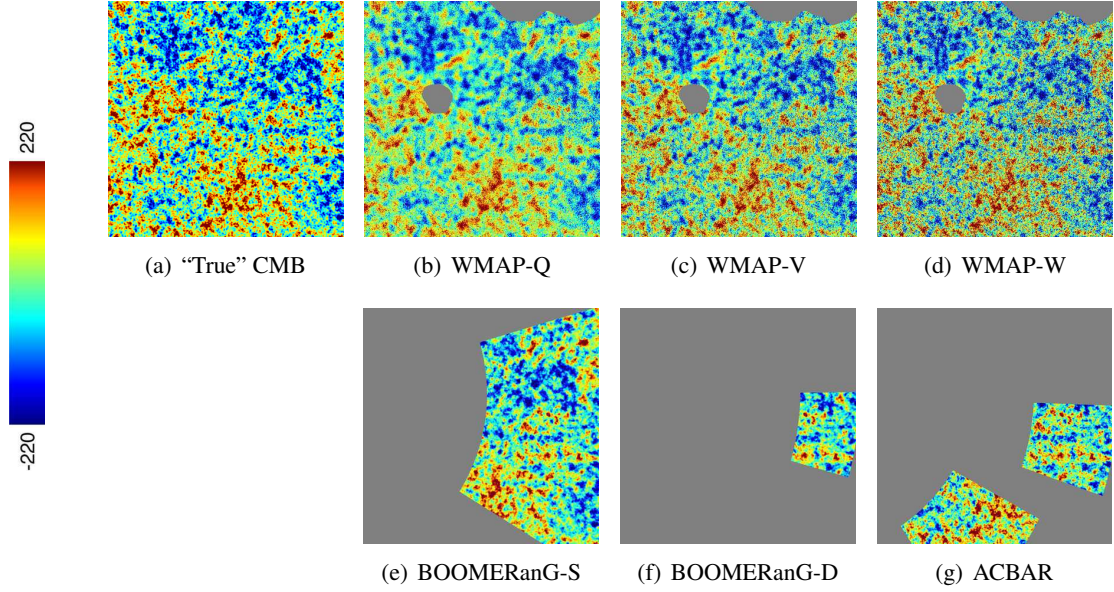


FIG. 6.9 – Simulated observations from model (6.3) for the 6 experiments described in Table 6.2, in a small patch around point  $(-40, -90)$ . The approximate size of the patch is  $38 \times 38$  degrees.

Key elements for this numerical experiment are illustrated in Fig. 6.9-6.10-6.11, where we have displayed respectively one random outcome of each experiment according to those simplified models, the maps of local noise levels and power spectra of the CMB and of the experiment's noise.

Fig. 6.13 displays the maps of the weights  $\omega_{k,e}^{(26)}$  (the 26th band is the multipole range  $700 < \ell \leq 800$ ). According to Eq. (6.23), all those weights belong to  $[0, 1]$  and for any fixed position, the sum of the weights over the six experiments is equal to one. Red regions indicate needlet coefficients which are far better observed in an experiment than in all others. Blue, light blue and orange regions are increasing but moderately low weights, showing that outside the small patches of BOOMERanG and ACBAR, most information on band 26 is provided by the channel W of WMAP. On the patches, needlet coefficients from WMAP are numerically neglected in the combination (6.21).

The debiased, squared, aggregated coefficients  $(\tilde{\gamma}_k^{(26)})^2 - (\tilde{n}_k^{(26)})^2$  are displayed on the left map from Fig. 6.14. All those coefficients are approximately unbiased estimators of  $C^{(26)}$ . The map of weights  $\tilde{w}_k^{(26)}$  is displayed on the right of Figure 6.14. More weight is given to regions which are covered by lower noise experiments. The final estimate is obtained by averaging the pixelwise multiplication of these two maps.

Figures 6.15 and 6.16 show the benefit of the aggregation of different experiments, in comparison with separate estimations. In CMB literature, error bars from different experiments are usually plotted on a same graph with different colors. For easier reading, we plot the output of single experiment NSE in separate panels (a,b and c). Panel (d) shows the output of the aggregated NSE, which improves the best single experiment uniformly over the frequency range, thanks to the locally adaptive combination of informations from all experiments.

Figure 6.12 highlights the cross-correlation between single experiment estimators and the final aggregated estimator. It provides a complementary insight on the relative weight of each experiment in

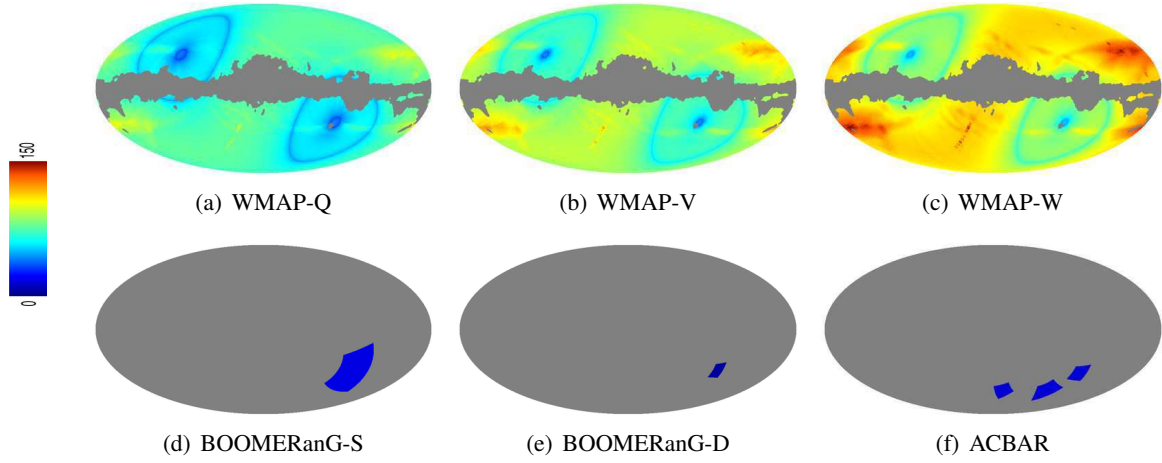


FIG. 6.10 – Coverage and local pixel noise levels of the six simple experiments described in Table 6.2.

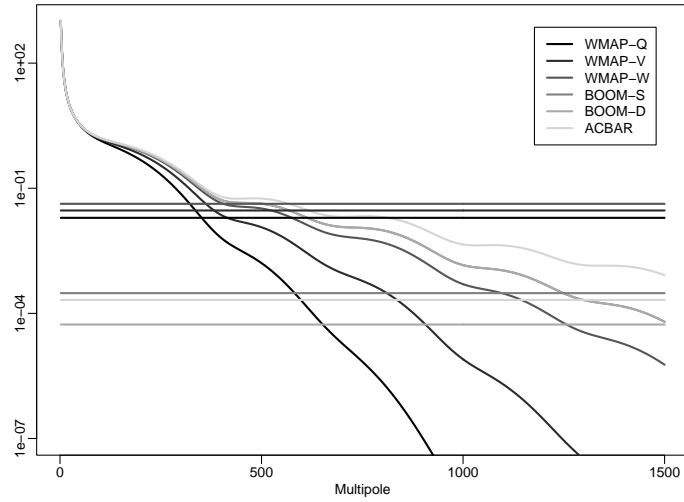
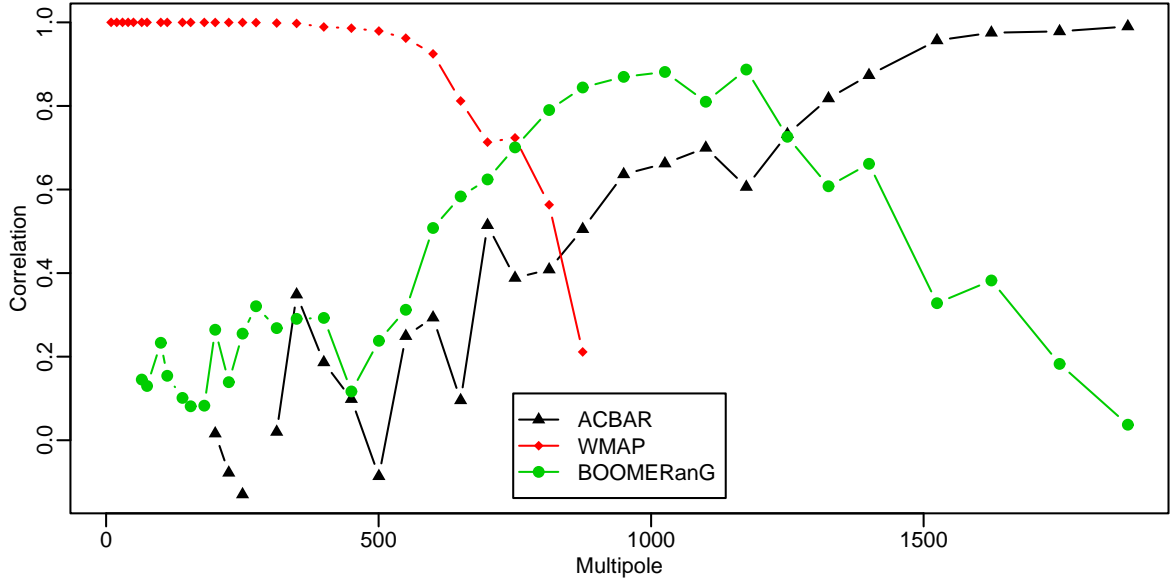
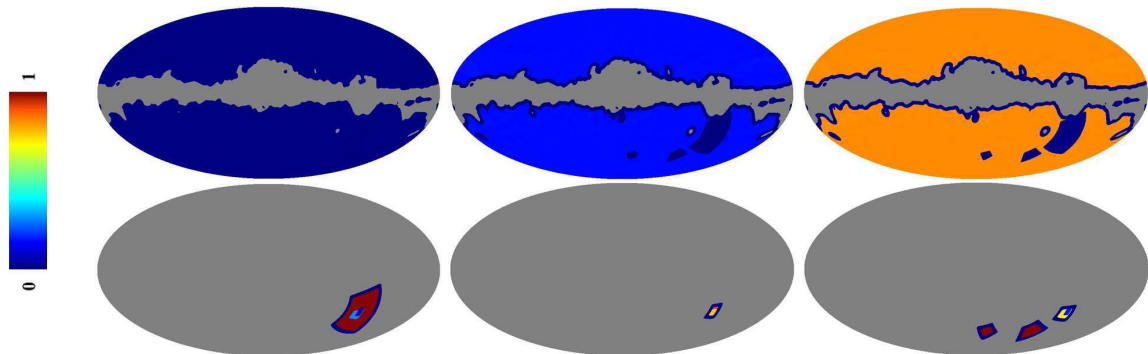


FIG. 6.11 – Spectra of the beamed CMB (with the BOOMERanG lines overplotted) and noise levels (horizontal lines) seen by the six experiments, as if they were full sky (the  $f_{\text{sky}}$  effect is not taken into account).



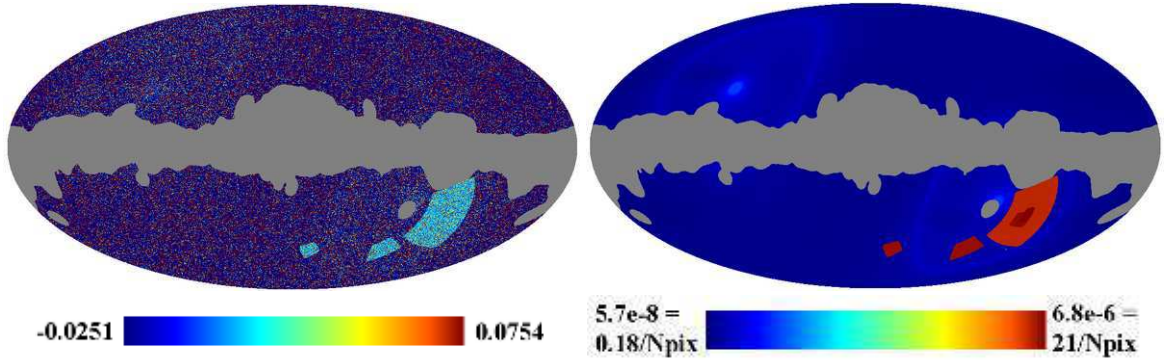
F . 6.12 – Correlation between the aggregated estimator and single experiments estimators. This provides insight on the contribution of each experiment into the final aggregated single spectral estimate.

the spectral domain. The WMAP-like measures are decisive for lower bands, whereas BOOMERanG and ACBAR ones give estimators very much correlated to the aggregated one at higher bands. The aggregated NSE is eventually almost identical to the estimator obtained from ACBAR alone.

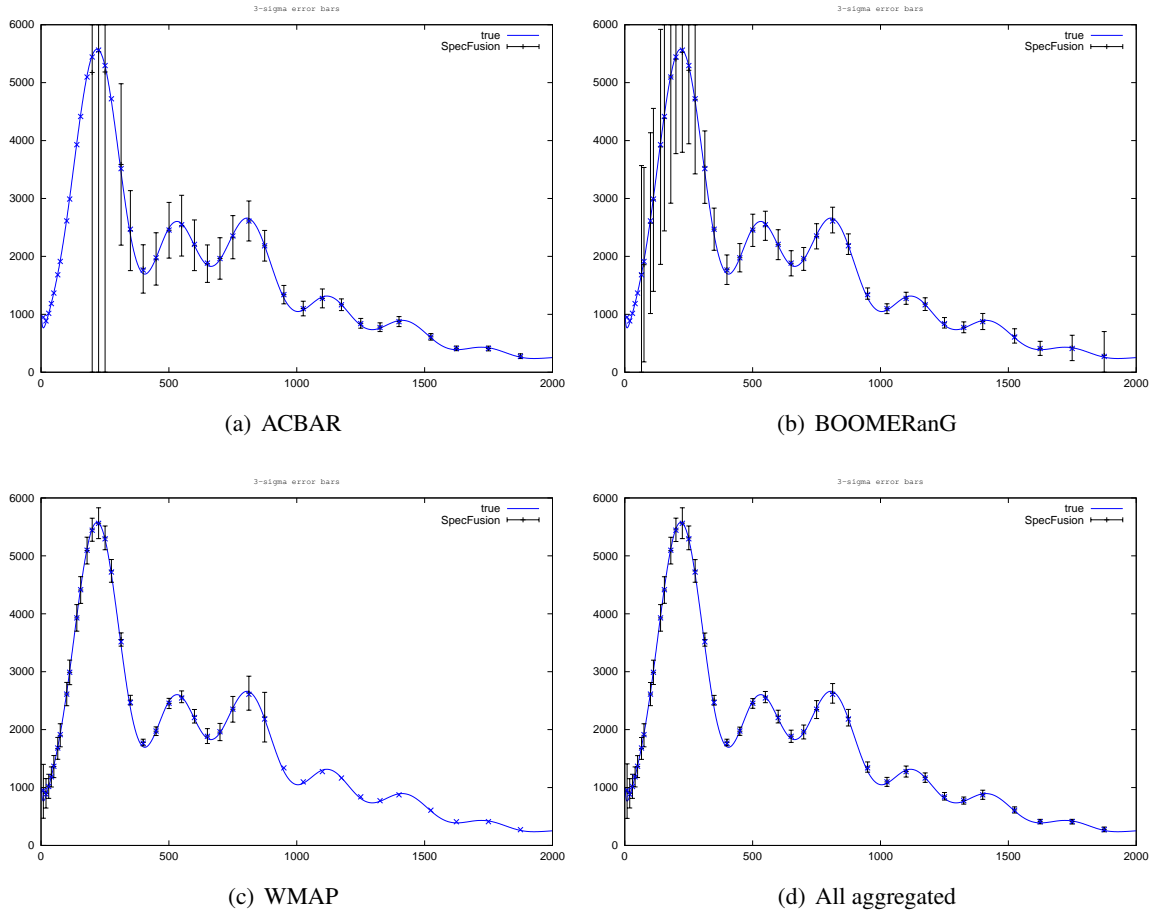


F . 6.13 – Method for aggregating experiments : Weights  $\omega_{k,e}^{(26)}$  for combining the needlet coefficients from the 26th band ( $700 < \ell \leq 800$ ) and the six experiments. From left to right and top to bottom : WMAP-Q, WMAP-V, WMAP-W, BOOMERanG-S, BOOMERanG-D and ACBAR.





F . 6.14 – Method for aggregating experiments : On the left : map of debiased squares of aggregated needlet coefficients, in the 26th band ( $700 < \ell \leq 800$ ). On the right : map of the weights  $w_{j,k}$  affected to those coefficients to estimate the power spectrum.



F . 6.15 – Results for the aggregated NSE. Error bars are estimated by 100 Monte Carlo simulations. The ACBAR power spectrum is computed using the single map needlet estimator described in Section 6.2.1, whereas the BOOMERanG and WMAP spectra are obtained using the aggregation of needlets coefficients from the two (BOOM-S and BOOM-D) and three (WMAP Q,V,W) maps respectively. The final spectrum (d) is obtained by aggregating all available needlet coefficients from the six maps.

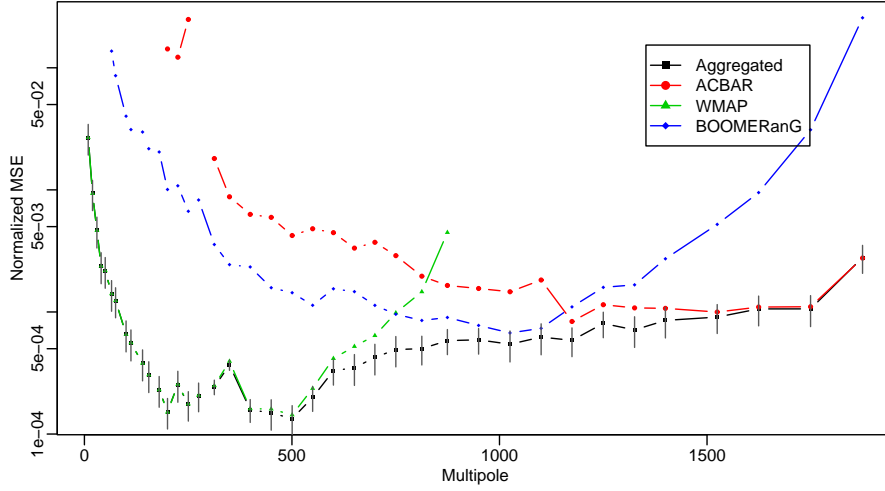


FIG. 6.16 – Mean-square error of the three single experiment NSE estimators and of the aggregated NSE estimator. The 2-sigma error bars reflect the imprecision in the Monte Carlo estimation of the MSE of the aggregated NSE. Up to those uncertainties, the aggregated estimator is uniformly better than the best of all experiments. The improvement is decisive in “crossing” regions, where two experiments perform comparably. The normalized MSE here is  $\mathbf{E}(\hat{C}^{(j)} - C^{(j)})^2 / (C^{(j)})^2$ .

## 6.4 Discussion

### Complexity

According to (6.5), the calculation of all the needlet coefficients takes one SHT and  $j_{\max}$  inverse SHT, where  $j_{\max}$  is the number of bands. The weights  $w_{j,k}$ ,  $\tilde{w}_{j,k}$  and  $\omega_{j,k,e}$  are obtained using simple operations on maps, so that the overall cost of the (aggregated) NSE scales as  $N^{3/2}$  operations. This is comparable to the cost of the PCL methods.

### Sensitivity to the noise knowledge

To be unbiased, the above described estimators require a perfect knowledge of the noise characteristics, as do Pseudo- $C_\ell$  estimators. In both cases, the uncertainty on the noise can be tackled using cross-spectrum, that removes the noise on the average provided that the noises from each experiment are independent. Indeed, for any pixel  $k$  far enough from the masks of experiment  $e$  and  $e'$ ,  $e \neq e'$ , we have

$$\mathbf{E}[\gamma_{j,k,e} \gamma_{j,k,e'}] = B_{j,e} B_{j,e'} C^{(j)}.$$

Thus, an unbiased spectral estimator is given by

$$\widehat{C^{(j)}}_{\text{cross}} = \sum_{k \in K_j} w_{j,k} \sum_{e \neq e'} (B_{j,e} B_{j,e'})^{-1} \gamma_{j,k,e} \gamma_{j,k,e'} \quad (6.29)$$

where the weights  $w_{j,k}$  depend on a preliminary estimate of the spectrum and a possibly imprecise estimate of the local and aggregated noise levels that enter in the variance of  $\gamma_{j,k,e} \gamma_{j,k,e'}$ . This has not

been investigated numerically yet but we can conjecture the qualitative results of this approach : more robustness with respect to noise misspecification but greater error bars than the NSE with perfectly known noise levels. Moreover, adapting the procedure described in Polenta *et al.* [122], one can test for noise misspecification, and for the correct removal of the noise by considering the difference between the NSE  $\widehat{C}^{(j)}$  and cross-spectrum NSE  $\widehat{C}^{(j)}_{\text{cross}}$ .

## 6.5 Conclusion

We have presented some potentialities of the needlets on the sphere for the angular power spectrum estimation. This tool is versatile and allows to treat consistently the estimation from a single map or from multiple maps. There remains many ways of improving or modify the method described in Section 6.2.

In the future, it is likely that again complementary data sets will co-exist. This is the case, in particular, for polarisation, for which Planck will measure the large scale CMB power on large scales with moderate sensitivity, while ground-based experiments will measure very accurately polarisation on smaller scales. Extensions to polarisation of the approach presented hers will likely be important for the best exploitation of such observations.

## 6.6 Appendices

### 6.6.1 Pseudo- $C_\ell$ estimators

Let  $T$  be a stationary process with power spectrum  $(C_\ell)_{\ell \geq 0}$ ,  $\mathcal{W}$  an arbitrary weight function (or mask) and

$$\widetilde{C}_\ell(\mathcal{W}) = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |\langle Y_{\ell m}, \mathcal{W}T \rangle|^2$$

the so-called *pseudo-power spectrum* of  $T$  with mask  $\mathcal{W}$ . The ensemble-average of this quantity is related to the true power spectrum by the formula

$$\mathbf{E}(\widetilde{C}_\ell) = \mathcal{M}_{\ell\ell'}(\mathcal{W})C_{\ell'}$$

where  $\mathcal{M}_{\ell\ell'}(\mathcal{W})$  is the doubly-infinite coupling matrix associated with  $\mathcal{W}$ , see [119, 89]. If  $U$  is a unit variance white pixel noise, denote by  $V_\ell \equiv 4\pi\sigma^2/N_{\text{pix}}$  its “spectrum” (see Appendix 6.6.2). Consider now the model  $X = \mathcal{W}_1 T + \mathcal{W}_2 U$ . Then, if  $\mathcal{M}_{\ell\ell'}(\mathcal{W}_1)$  is full-rank,

$$(\mathcal{M}_{\ell\ell'}(\mathcal{W}_1))^{-1} \left\{ \widetilde{C}_\ell(\mathcal{W}_1) - \mathcal{M}_{\ell\ell'}(\mathcal{W}_2)V_{\ell'} \right\}$$

is an unbiased estimator of  $C_{\ell'}$ . It is obtained by deconvolving and debiasing the empirical spectrum. The observation model (6.2) with no beam coincides with the preceding framework with  $\mathcal{W}_1 = \mathcal{W}$  and  $\mathcal{W}_2 = \sigma\mathcal{W}$ . This leads to the *uniform-weights pseudo- $C_\ell$  estimator* (PCLU). One can also divide all the observations by  $\sigma^2$ , yielding to a similar scheme with  $\mathcal{W}_1 = \sigma^{-2}\mathcal{W}$  and  $\mathcal{W}_2 = \sigma^{-1}\mathcal{W}$ . This is the *variance-weighted pseudo- $C_\ell$  estimator* (PCLW). Both are used by the WMAP collaboration [86]. The uniform weights lead to better estimates in the high SNR regime (low  $\ell$ 's) whereas the flat weights

perform better at low SNR (high  $\ell$ 's). Efsthathiou [74] showed that the Pseudo- $C_\ell$  estimator is statistically equivalent to the maximum likelihood estimator asymptotically as  $\ell$  goes to infinity. He also proposed an implementation of a smooth transition between those two regimes.

### 6.6.2 What “noise spectrum” means

Let  $\nu$  denote the noise. It is defined on pixels and supposed centered, Gaussian, independent from pixel to pixel, and of variance  $\sigma^2(\xi)$ , *i.e.*

$$\nu_k = \sigma(\xi_k)U_k, \quad k = 1, \dots, N,$$

with  $U_1, \dots, U_N \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ . Define  $\nu_{\ell,m} := \sum_k \lambda_k \nu_k Y_{\ell,m}(\xi_k)$ , and call them (abusively) the “discretized” multipole moments of the noise, which do not have any continuous counterpart because  $\nu$  is not defined on the whole sphere. Define the corresponding discretized empirical spectrum  $\bar{N}_\ell := \frac{1}{2\ell+1} \sum_m \nu_{\ell,m}^2$ , then

$$\begin{aligned} \mathbf{E}(\nu_{\ell,m} \nu_{\ell',m'}) &= \sum_k \lambda_k^2 \sigma^2(\xi_k) Y_{\ell,m}(\xi_k) Y_{\ell',m'}(\xi_k) \\ \bar{N}_\ell &= \frac{1}{2\ell+1} \sum_{k,k'} \lambda_k \lambda_{k'} \sigma(\xi_k) \sigma(\xi_{k'}) U_k U_{k'} L_\ell(\langle \xi_k, \xi_{k'} \rangle) \\ \text{and } \mathbf{E}(\bar{N}_\ell) &= \frac{1}{4\pi} \sum_k \lambda_k^2 \sigma^2(\xi_k) =: N_\ell \end{aligned}$$

This sequence  $N_\ell$  can be thought of as the pixel-noise spectrum. Note that if  $\lambda_k = \frac{4\pi}{N_{\text{pix}}}$ ,  $k = 1, \dots, N$ , then  $N_\ell = \frac{1}{N_{\text{pix}}} \int \sigma^2(\xi) d\xi$ . If the noise is moreover homogeneous,  $\sigma(\xi) \equiv \sigma$ , then  $\mathbf{E}(\nu_{\ell,m} \nu_{\ell',m'}) = \frac{4\pi\sigma^2}{N_{\text{pix}}} \delta_{\ell,\ell'} \delta_{m,m'}$ .

### 6.6.3 Variance estimation by aggregation of experiments with independent heteroscedastic noise

Consider the model

$$Y_{k,e} = X_k + n_{k,e} Z_{k,e}$$

where  $\mathbf{X} := [X_k]_{k \in [1,N]}$  and  $\mathbf{Z} := [Z_{k,e}]_{(k,e) \in [1,N] \times [1,E]}$  are independent,  $X_k \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, C)$ ,  $Z_{k,e} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$  and the noise standard deviations  $n_{k,e}$  are known. This corresponds to the observation of the same signal  $\mathbf{X}$  by  $E$  independent experiments, the observations being tainted by independent but heteroscedastic errors. Let  $\mathbf{Y}_k := [Y_{k,e}]_{e \in [1,E]}$  be the vector of observations at point (or index in a general framework)  $k$ , and let  $\mathbf{Y} := ([\mathbf{Y}_k^T]_{k \in [1,N]})^T$  be the full vector of observations. The covariance matrix of  $\mathbf{Y}_k$  is  $\mathbf{R}_k := \mathbf{1}\mathbf{1}^T C + \mathbf{N}_k$  where  $\mathbf{N}_k := \text{diag}(n_{k,e}^2)_{e \in [1,E]}$  and  $\mathbf{1}$  is the  $E \times 1$  vector of ones. By independence of the  $\mathbf{Y}_k$ 's, the negative log-likelihood of  $C$  given  $\mathbf{Y}$  thus writes

$$\begin{aligned} \mathcal{L}(C) &:= -2 \log(P(\mathbf{Y}|C)) = -2 \sum_k \log(P(\mathbf{Y}_k|C)) \\ &= \sum_k \mathbf{Y}_k^T \mathbf{R}_k^{-1} \mathbf{Y}_k + \log \det \mathbf{R}_k. \end{aligned}$$

Denote

$$\tilde{n}_k := (\mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{1})^{-1/2} = \left( \sum_e (n_{k,e})^{-2} \right)^{-1/2}. \quad (6.30)$$

It is immediate to check the following identity which will be used below :

$$\mathbf{R}_k^{-1} \mathbf{1} = \frac{\mathbf{N}_k^{-1} \mathbf{1}}{1 + C \tilde{n}_k^2}.$$

Define  $\widehat{\mathbf{R}}_k := \mathbf{Y}_k \mathbf{Y}_k^T$ . The derivative of the negative log-likelihood writes

$$\begin{aligned} \mathcal{L}'(C) &= \sum_k -\mathbf{Y}_k^T \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial C} \mathbf{R}_k^{-1} \mathbf{Y}_k + \text{tr} \left( \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial C} \right) \\ &= \sum_k \text{tr} \left( -\mathbf{Y}_k^T \mathbf{R}_k^{-1} \mathbf{1} \mathbf{1}^T \mathbf{R}_k^{-1} \mathbf{Y}_k \right) + \text{tr} \left( \mathbf{R}_k^{-1} \mathbf{1} \mathbf{1}^T \right) \\ &= \sum_k \mathbf{1}^T \mathbf{R}_k^{-1} (\mathbf{R}_k - \widehat{\mathbf{R}}_k) \mathbf{R}_k^{-1} \mathbf{1} \\ &= \sum_k \frac{\mathbf{1}^T \mathbf{N}_k^{-1} (\mathbf{R}_k - \widehat{\mathbf{R}}_k) \mathbf{N}_k^{-1} \mathbf{1}}{(1 + C \tilde{n}_k^2)^2} \\ &= C \sum_k \frac{(\mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{1})^2}{(1 + C \tilde{n}_k^2)^2} - \sum_k \frac{(\mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{Y}_k)^2 - \mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{1}}{(1 + C \tilde{n}_k^2)^2} \\ &= C \sum_k (C + \tilde{n}_k^2)^{-2} - \sum_k \frac{(\tilde{n}_k^2 \mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{Y}_k)^2 - \tilde{n}_k^2}{(C + \tilde{n}_k^2)^2}. \end{aligned}$$

It follows that the likelihood is maximized for

$$C = \widehat{C}(w) := \sum_k w_k(C, \mathbf{N}) \left[ (\tilde{n}_k^2 \mathbf{1}^T \mathbf{N}_k^{-1} \mathbf{Y}_k)^2 - \tilde{n}_k^2 \right]$$

with

$$w_k(C, \mathbf{N}) := (C + \tilde{n}_k^2)^{-2} \left[ \sum_i (C + \tilde{n}_i^2)^{-2} \right]^{-1}. \quad (6.31)$$

As the optimal weights depend on  $C$ , this only defines implicitly the ML estimator. For some approximate spectrum  $C^0$ , the proposed explicit NSE is given by  $\widehat{C}(\widehat{w}_k)$  with  $\widehat{w}_k = w_k(C^0, \mathbf{N})$ .

### Particular case of a single experiment

In the particular case of a single experiment ( $E = 1$ ) with heteroscedastic noise, following the model

$$Y_k = X_k + n_k Z_k,$$

the likelihood is maximized for

$$C = \widehat{C}(w) := \sum_k w_k(C, \mathbf{N}) (Y_k^2 - n_k^2) \quad (6.32)$$

with  $w_k(C, \mathbf{N})$  defined by Eq. (6.31), and again, assuming that  $w_k$  is poorly sensitive to  $C$ , the NSE is  $\widehat{C}(\widehat{w}_k(C^0))$  for some approximate spectrum  $C^0$ .

## **Acknowledgements**

The ADAMIS team at APC has been partly supported by the Astro-Map and Cosmostat ACI grants of the French ministry of research, for the development of innovative CMB data analysis methods. The results in this paper have been derived using the HEALPix package [84]. Our pipeline is mostly implemented in octave ([www.octave.org](http://www.octave.org)).

## Chapitre 7

# Spectral estimation on the sphere with needlets : high frequency asymptotics

Ce chapitre reproduit l'article du même nom, G. Faÿ, F. Guillaux (soumis). Seules certaines notations ont été modifiées pour la cohérence du manuscrit.

**Abstract.** The angular power spectrum of a stationary random field on the sphere is estimated from the needlet coefficients of a single realization, observed with increasingly fine resolution. The estimator we consider is similar to the one recently used in practice by (Faÿ et al. 2008) to estimate the power spectrum of the Cosmic Microwave Background. The consistency of the estimator, in the asymptotics of high frequencies, is proved for a model with a stationary Gaussian field corrupted by heteroscedastic noise and missing data.

### Sommaire

7.1	Introduction . . . . .	131
7.2	Model and settings . . . . .	132
7.2.1	Gaussian stationary spherical fields . . . . .	132
7.2.2	Sampling on the sphere . . . . .	132
7.2.3	Observation model . . . . .	133
7.2.4	Needlets and statistical properties of needlet coefficients . . . . .	134
7.3	Estimation results . . . . .	136
7.3.1	Parameters and estimators . . . . .	137
7.3.2	Consistency of the estimators . . . . .	138
7.4	Simulations . . . . .	139
7.4.1	Settings . . . . .	140
7.4.2	Results . . . . .	142
7.5	Conclusion . . . . .	142
7.6	Proofs . . . . .	143
7.6.1	Proof of Propositions 7.1 and 7.2 . . . . .	143
7.6.2	Proof of Proposition 7.5 . . . . .	144
7.6.3	Proof of Proposition 7.6 . . . . .	145
7.6.4	Proof of Theorems 7.1 and 7.2 . . . . .	146
7.6.5	End of proof of Theorem 7.1 . . . . .	147

## 7.1 Introduction

In many application domains (geophysics, cosmology, hydrodynamics, computer vision, etc.), data are defined on the sphere. If the data fit the model of a stationary stochastic field, their second order characteristics, summarized by the angular power spectrum, is of great importance. It contains all the distribution information in the case of a Gaussian stationary process. Ordinary spherical harmonic transform (SHT), the equivalent of the Fourier Series on the circle, provides a fast and efficient method for spectral estimation in the idealistic case of a fully and perfectly observed sphere. We are mainly concerned here with the asymptotics of high angular frequencies.

However, rarely the data are available on the whole sphere. Often it is observed under non-stationary contaminants. This is the case for the cosmic microwave background (CMB) which is a major motivation for this work. For those reasons, during the past decade, *localized* analysis for spherical data has motivated many developments ; see [184, 51, 15, 134, 28] and the references therein.

The wavelets provide a powerful framework for dealing with non-stationarities. A recent construction of wavelet frames (needlets) by [18, 33] has proved efficient to analyze stationary spherical processes, thanks to their good localization property. The statistical properties of these frames' coefficients, in a noiseless context, are established in [151, 149, 171]. In [151], a Central Limit theorem is proved for functionals of the needlet coefficients of a pure, noiseless, unmasked, Gaussian stationary field. This is supplemented in [149] with bootstrap estimators of the variance of such statistics ; the authors show that these bootstrap estimators are “consistent” (in a high frequency sense ; see below) and that the CLT holds when the true variance is replaced by its bootstrap estimate.

In the time series literature, wavelets are used for spectral estimation whether in a semi-parametric (see e.g. [176]) or a non-parametric [154] context. Our observation model, in addition to being spherical, has the particularity of presenting quite general non stationarity (in the structure of the noise) and we failed to find any reference on the subject even for processes living on Euclidean spaces.

In this paper, we establish the consistency of a spectral estimator constructed on the needlets coefficients in high-frequency asymptotics.

The paper is organized as follows. In Section 7.2 we present the model, including assumptions on the way the process is sampled. In Section 7.3 we define the needlet spectral estimators and state the consistency results that hold true under realistic conditions. The finite sample behavior of the estimator is explored by numerical simulations in Section 7.4. Our conclusion is given in Section 7.5 and proofs are postponed to Section 7.6.



## 7.2 Model and settings

### 7.2.1 Gaussian stationary spherical fields

Consider the unit sphere  $\mathbb{S}^2$  in  $\mathbb{R}^3$  with generic element  $\xi$ . The geodesic distance is given by  $d(\xi, \xi') := \arccos(\xi \cdot \xi')$  where  $\xi \cdot \xi'$  denotes the usual dot product between  $\xi$  and  $\xi'$  (considered as vectors in  $\mathbb{R}^3$ ). The uniform measure  $d\xi$  is the unique positive measure on  $\mathbb{S}^2$  which is invariant by rotation, with total mass  $4\pi$ . Let  $\mathbb{L}^2(\mathbb{S}^2) := \mathbb{L}^2(\mathbb{S}^2, d\xi)$  be the Hilbert space of complex-valued square integrable functions. We have the following decomposition :  $\mathbb{L}^2(\mathbb{S}^2) = \bigoplus_{\ell=0}^{\infty} \mathbb{H}_{\ell}$  where  $\mathbb{H}_{\ell}$  is the vector space of spherical harmonics of degree  $\ell$ , *i.e.* restrictions to the sphere of homogeneous polynomials of degree  $\ell$  in  $\mathbb{R}^3$  which are harmonic (or, equivalently, the restriction of which are eigenvectors of the spherical Laplacian with eigenvalues  $\ell(\ell+1)$ ). The usual spherical harmonics  $\mathcal{Y}_{\ell,m}(\xi)$ ,  $-\ell \leq m \leq \ell$ , constitute an orthonormal basis of  $\mathbb{H}_{\ell}$ . Therefore, the set of all spherical harmonics  $\mathcal{Y}_{\ell,m}$ ,  $\ell \geq 0$ ,  $-\ell \leq m \leq \ell$ , is an orthonormal basis of  $\mathbb{L}^2(\mathbb{S}^2)$ .

In this paper, we shall be concerned with a zero-mean, mean square continuous and real-valued random field  $X(\xi)$ . We shall assume that  $X$  is second-order stationary, that is  $\mathbf{E}[X(\rho\xi)X(\rho\xi')] = \mathbf{E}[X(\xi)X(\xi')]$  for all  $\rho \in SO(3)$ . Then the spherical harmonics coefficients of  $X$ ,  $a_{\ell,m} := \langle X, \mathcal{Y}_{\ell,m} \rangle$ , are square integrable random variables which verify  $\mathbf{E}[a_{\ell,m}a_{\ell',m'}^*] = \delta_{\ell,\ell'}\delta_{m,m'}C_{\ell}$  for  $m, m' \geq 0$  and  $a_{\ell,-m} = a_{\ell,m}^*$ . The inverse spherical harmonics transform reads :  $X(\xi) = \sum_{\ell \geq 0} \sum_{m=-\ell}^{\ell} a_{\ell,m} \mathcal{Y}_{\ell,m}(\xi)$ . The last equality holds in the sense that  $\mathbf{E} \int_{\mathbb{S}^2} \left( X(\xi) - \sum_{\ell=0}^{\ell^{max}} \sum_{m=-\ell}^{\ell} a_{\ell,m} \mathcal{Y}_{\ell,m}(\xi) \right)^2 d\xi \xrightarrow{\ell^{max} \rightarrow \infty} 0$ . The sequence  $(C_{\ell})_{\ell \geq 0}$  is called the (angular) power spectrum of  $X$ . Let  $L_{\ell}$  denote the Legendre polynomial of degree  $\ell$  normalized by  $L_{\ell}(1) = \frac{2\ell+1}{4\pi}$ . The angular power spectrum is linked to the angular correlation of  $X$  by the relation  $\sum_{\ell \geq 0} C_{\ell} L_{\ell}(\cos \theta) = \mathbf{E}[X(\xi)X(\xi')]$  for all pairs of points such that  $d(\xi, \xi') = \theta$ . The square integrability of  $X$  is equivalent to the condition  $\sum_{\ell \geq 0} (2\ell+1)C_{\ell} < \infty$ .

We shall also assume that  $X$  is Gaussian. This additional assumption is known to be true if and only if the coefficients  $a_{\ell,m}$ ,  $\ell \geq 0$ ,  $m \geq 0$  are independent (see [153] for the “only if” part). As mentioned in the Introduction, the finite-dimensional distributions of a Gaussian stationary field are entirely determined by the second-order characteristics, that is by the angular power spectrum of the field. For instance, the second-order stationarity is equivalent, under Gaussian assumption, to the strict stationarity, *i.e.* for all  $\rho \in SO(3)$  and  $\xi_1, \dots, \xi_n \in \mathbb{S}^2$  the two vectors  $(X(\rho\xi_1), \dots, X(\rho\xi_n))$  and  $(X(\xi_1), \dots, X(\xi_n))$  have the same distribution.

### 7.2.2 Sampling on the sphere

In any real-life situation, only discretized versions of  $X$  are available, and consequently spherical harmonic coefficients are exactly computable only if  $X$  is  $\ell^{max}$ -band-limited, that is all the  $a_{\ell,m} = 0$ ,  $\ell > \ell^{max}$ , for some  $\ell^{max}$  which depends essentially on the number of observed points. The discretization of the sphere and achievement of cubature formulas for geodetic functions is a non-trivial task. During the last decade it was shown ([29, 33]) that there exists a constant  $c_0 > 0$  such that for all  $\ell^{max} \in \mathbb{N}^*$  there exists a set  $(\xi_k, \lambda_k)_{k \in \{1, \dots, N\}} \in (\mathbb{S}^2 \times \mathbb{R}_+^*)^N$  of cubature points and weights (referred to as a

pixelization of order  $L$ ) with the following properties.

$$\text{For all } f \in \bigoplus_{\ell=0}^{\ell^{\max}} \mathbb{H}_\ell, \quad \int_{\mathbb{S}^2} f(\xi) d\xi = \sum_{k=1}^N \lambda_k f(\xi_k) \quad (\text{cubature formula}). \quad (7.1a)$$

$$c_0^{-1} (\ell^{\max})^2 \leq N \leq c_0 (\ell^{\max})^2. \quad (7.1b)$$

$$c_0^{-1} (\ell^{\max})^{-2} \leq \min_{1 \leq k \leq N} \lambda_k \leq \max_{1 \leq k \leq N} \lambda_k \leq c_0 (\ell^{\max})^{-2}. \quad (7.1c)$$

$$c_0^{-1} (\ell^{\max})^{-1} \leq \sup_{\xi \in \mathbb{S}^2} d(\xi, \{\xi_k\}_{k \in \{1, \dots, N\}}) \leq c_0 (\ell^{\max})^{-1}. \quad (7.1d)$$

$$c_0^{-1} (\ell^{\max})^{-1} \leq \min_{1 \leq k < k' \leq N} d(\xi_k, \xi_{k'}) \leq c_0 (\ell^{\max})^{-1}. \quad (7.1e)$$

The following two lemmas derive straightforwardly from these pixelization properties. The first one is proved in [151] and the second one follows from a simple covering argument (see [149, Lemma 5] for a similar statement).

**Lemma 7.1** ([151]). *For all  $M \geq 3$ , there exists a constant  $c > 0$  such that for all  $\ell^{\max} \in \mathbb{N}^*$ ,  $\xi \in \mathbb{S}^2$  and for all pixelization of order  $\ell^{\max}$  we have*

$$\sum_{k=1}^N \frac{1}{(1 + \ell^{\max} d(\xi, \xi_k))^M} \leq c.$$

**Lemma 7.2.** *There exists a constant  $c > 0$  such that for all  $\ell^{\max} \in \mathbb{N}^*$ ,  $\xi \in \mathbb{S}^2$  and  $\delta > c_0 (\ell^{\max})^{-1}$  and for all pixelization of order  $\ell^{\max}$  we have*

$$c^{-1} \delta^2 (\ell^{\max})^2 \leq \text{Card} \{k \in \{1, \dots, N\} : d(\xi, \xi_k) \leq \delta\} \leq c \delta^2 (\ell^{\max})^2.$$

### 7.2.3 Observation model

We are now in position to give a description of our statistical model. Assume that we observe a noisy and sampled version of low-passed  $X$  at successive *scales*  $j$ , with some missing (or attenuated) data.

More precisely, for every  $j \in \mathbb{N}$ , given some  $\ell_j^{\max} \in \mathbb{N}^*$ ,

- let  $(\xi_{j,k}, \lambda_{j,k})_{k \in \{1, \dots, N_j\}}$  be a pixelization of order  $4\ell_j^{\max}$  ;
- let  $W_j : k \in \{1, \dots, N_j\} \mapsto W_{j,k} \in [0, 1]$  and  $\sigma_j : k \in \{1, \dots, N_j\} \mapsto \sigma_{j,k} \in \mathbb{R}_+$  be deterministic, known, functions ;
- and let  $B_j : \ell \in \mathbb{N} \mapsto B_{j,\ell} \in \mathbb{R}$  such that  $B_{j,\ell} = 1$  if  $\ell \leq \ell_j^{\max}$  and  $B_{j,\ell} = 0$  if  $\ell > 2\ell_j^{\max}$ .

We observe

$$Y_{j,k} := W_{j,k} [X_j(\xi_{j,k}) + Z_{j,k}], \quad j \in \mathbb{N}, k \in \{1, \dots, N_j\} \quad (7.2)$$

where  $Z_{j,k} := \sigma_{j,k} U_{j,k}$  and  $U_{j,k}$ ,  $j \in \mathbb{N}$ ,  $k \in \{1, \dots, N_j\}$  is a triangular array of standard and independent Gaussian random variables and independent of the process  $X$ . The process  $X_j$  is defined by

$$X_j := \sum_{\ell \geq 0} \sum_{m=-\ell}^{\ell} B_{j,\ell} a_{\ell,m} \mathcal{Y}_{\ell,m}, \quad a_{\ell,m} = \langle X, \mathcal{Y}_{\ell,m} \rangle. \quad (7.3)$$

In any CMB experiment, some smoothing is induced by the instrumental beam. Eq (7.3) is a idealistic version of this low-pass operation.

Without loss of generality,  $j \mapsto \ell_j^{\max}$  is supposed non-decreasing. In the following, we call  $W_j$  the *mask*. The particular case of  $W_j$  taking its values in  $\{0, 1\}$  corresponds to missing data.

In other words, a single realization of  $X$  is considered, but independent and noisy measures with an increasing spatial resolution are available. This corresponds, for instance, to the observation model of the CMB. The latter is modeled by astrophysicists as the single realization of a stationary Gaussian process. Its observation is achieved by more and more precise instruments, involving their own observation noise, sky coverage and instrumental beam. Full sky map of moderate resolutions (*e.g.* maps provided by the WMAP collaboration [53]) and observations of small and clean patches of the sky at very high resolution (*e.g.* maps from ACBAR experiment [124]) are available simultaneously. Cosmologists aggregate information for those maps to give a large band estimation of the angular power spectrum.

## 7.2.4 Needlets and statistical properties of needlet coefficients

### General framework

The needlets are second-generation wavelet frames which were introduced by [33] (see also [18, 25]). Let us recall below their definition and first properties, the proofs of which are either referred to existing literature or postponed to Section 7.6.

Start from the fact that the orthogonal projection on  $\mathbb{H}_\ell$  has a kernel involving Legendre polynomials, namely

$$\forall f \in \mathbb{L}^2(\mathbb{S}^2), (\Pi_{\mathbb{H}_\ell} f)(\xi) := \sum_{m=-\ell}^{\ell} \langle f, \mathcal{Y}_{\ell,m} \rangle \mathcal{Y}_{\ell,m}(\xi) = \int_{\mathbb{S}^2} L_\ell(\xi \cdot \xi') f(\xi') d\xi'.$$

Instead of considering single frequencies  $\ell$ , we shall combine them within frequency bands. For this purpose, define a sequence of functions  $b_j : \ell \in \mathbb{N} \mapsto b_{j,\ell} \in \mathbb{R}$ ,  $j \in \mathbb{N}$ , called (frequency) *window functions*. The window  $b_j$  is supposed to be supported in  $[0, \tilde{\ell}_j^{\max}]$  for some  $\tilde{\ell}_j^{\max} \in \mathbb{N}$ . The kernels  $\Psi_j : (\xi, \xi') \mapsto \sum_{\ell \geq 0} b_{j,\ell} L_\ell(\xi \cdot \xi')$  and  $\Lambda_j : (\xi, \xi') \mapsto \sum_{\ell \geq 0} (b_{j,\ell})^2 L_\ell(\xi \cdot \xi')$  have the two following obvious properties. First, for all  $f \in \mathbb{L}^2(\mathbb{S}^2)$ ,  $f(\xi) = \sum_{j \in \mathbb{N}} \int_{\mathbb{S}^2} \Lambda_j(\xi, \xi') f(\xi') d\xi'$ . Second,  $\Lambda_j(\xi, \xi') = \int_{\mathbb{S}^2} \Psi_j(\xi, \xi'') \Psi_j(\xi'', \xi') d\xi''$ .

The discretization of the above kernels leads to the following spherical functions called *needlets*. For each scale  $j \in \mathbb{N}$ , given a pixelization  $(\xi_{j,k}, \lambda_{j,k})_{k \in \{1, \dots, N_j\}}$  of order at least  $2\tilde{\ell}_j^{\max}$ , define

$$\psi_{j,k}(\xi) := \sqrt{\lambda_{j,k}} \sum_{\ell \geq 0} b_{j,\ell} L_\ell(\xi \cdot \xi_k) \quad .$$

The needlets  $\psi_{j,k}$ ,  $j \in \mathbb{N}$ ,  $k \in \{1, \dots, N_j\}$  constitute a frame of  $\mathbb{L}^2(\mathbb{S}^2)$  [33, 25] if for all  $\ell \in \mathbb{N}$ ,  $\sum_{j \in \mathbb{N}} (b_{j,\ell})^2 = 1$ . For any (possibly random) function  $f$  in  $\mathbb{L}^2(\mathbb{S}^2)$ , the coefficients  $\langle f, \psi_{j,k} \rangle$  are renormalized for sake of notational simplicity : we shall handle the *needlet coefficients*

$$\gamma_{j,k} := (\lambda_{j,k})^{-1/2} \langle f, \psi_{j,k} \rangle \quad .$$

If  $f$  is  $\tilde{\ell}_j^{\max}$ -band-limited, one can compute practically the coefficients  $\gamma_{j,k}$ ,  $j \leq J$ , in the spherical harmonics domain, from the values of  $f$  on the cubature points, as made explicit by the following diagram.

$$(f(\xi_k))_{1 \leq k \leq N} \xrightarrow{\text{SHT}} (a_{\ell,m})_{\ell \leq \tilde{\ell}_j^{\max}} \xrightarrow{\times} (b_{j,\ell} a_{\ell,m})_{\ell \leq \tilde{\ell}_j^{\max}} \xrightarrow{\text{SHT}^{-1}} (\gamma_{j,k})_{1 \leq k \leq N_j} \quad (7.4)$$

The initial pixelization  $(\xi_k, \lambda_k)_{1 \leq k \leq N}$  must be of order at least  $2\tilde{\ell}_j^{\max}$ . SHT denotes spherical harmonics transform, computed from the samples of  $X$  and of the  $\mathcal{Y}_{\ell,m}$ 's thanks to (7.1a). Double arrows denotes  $J$  operations.

Since the needlet coefficients at a given scale  $j$  depend only on a finite number of values of the function  $f$ , it is possible to generalize this notion to an arbitrary (possibly random) finite sequence  $(f_k)_{1 \leq k \leq N_j} \in \mathbb{R}^{N_j}$ . The *needlet coefficients* of such sequence are the quantities

$$(\lambda_{j,k})^{-1} \sum_{p=1}^{N_j} \lambda_{j,p} \psi_{j,k}(\xi_{j,p}) f_p = \sum_{\ell \geq 0} \sum_{m=-\ell}^{\ell} b_{j,\ell} \mathcal{Y}_{\ell,m}(\xi_{j,k}) \sum_{p=1}^{N_j} \lambda_{j,p} \mathcal{Y}_{\ell,m}(\xi_{j,p}) f_p.$$

If  $f_k = f(\xi_{j,k})$  for some  $f \in \bigoplus_{\ell=0}^{\ell_j^{\max}}$  and  $\xi_k$  are the points of a pixelization of order at least  $2\tilde{\ell}_j^{\max}$ , then the above expressions are equal to  $\gamma_{j,k}$ .

Let us give the first properties of the needlet coefficients of a random field at a given scale  $j \in \mathbb{N}$ . Let  $X$  be a stationary field,  $X_j$  like in Eq. (7.3) with  $B_{j,\ell} = 1$  if  $\ell \leq \tilde{\ell}_j^{\max}$  and  $B_{j,\ell} = 0$  if  $\ell > 2\tilde{\ell}_j^{\max}$ , and  $(\xi_{j,k}, \lambda_{j,k})_{1 \leq k \leq N_j}$  a pixelization of order  $4\tilde{\ell}_j^{\max}$ . The needlet coefficients of  $X$  are denoted  $\eta_{j,k}$ . They are also the needlet coefficients of  $X_j$  since  $X$  and  $X_j$  have the same spherical harmonics coefficients up to the frequency  $\ell = \tilde{\ell}_j^{\max}$ . In the presence of an additive noise  $Z_{j,k}$  in the observation of  $X(\xi_{j,k})$ , the “observed” needlet coefficients computed by (7.4) from  $X+Z$  write  $\eta_{j,k} + \zeta_{j,k}$ , where the coefficients  $\zeta_{j,k}$  are the needlet coefficients of  $Z$ . The next results provide the covariance structure of those coefficients at scale  $j$ . In our model,  $X$  and  $Z$  are supposed Gaussian. In this case, their needlets coefficients are Gaussian too.

**Proposition 7.1** ([151]). *Denote  $(C_\ell)_{\ell \geq 0}$  the power spectrum of  $X$ . Its needlet coefficients  $\eta_{j,k}$  are centered, with covariances given by*

$$\text{Cov}[\eta_{j,k}, \eta_{j,k'}] = \sum_{\ell \geq 0} (b_{j,\ell})^2 C_\ell L_\ell(\xi_{j,k} \cdot \xi_{j,k'}).$$

**Proposition 7.2.** *Assume  $Z$  of the form  $Z_{j,k} = \sigma_{j,k} U_{j,k}$  where the  $U_{j,k}$  are uncorrelated, centered and unit variance random variables. The needlet coefficients  $\zeta_{j,k}$  of  $Z$  are centered, with covariances given by*

$$\text{Cov}[\zeta_{j,k}, \zeta_{j,k'}] = \frac{1}{\sqrt{\lambda_{j,k} \lambda_{j,k'}}} \sum_{p=1}^{N_j} (\lambda_{j,p} \sigma_{j,p})^2 \psi_{j,k}(\xi_{j,p}) \psi_{j,k'}(\xi_{j,p}).$$

### $\mathcal{B}$ -adic needlets

In this paper we shall fix some constant  $\mathcal{B} > 1$  and consider  $\mathcal{B}$ -adic window functions.

**Assumption 7.1.** *There exist  $M \geq 3$  and a  $M$ -differentiable real function  $\mathbf{a}$  supported in  $[-\mathcal{B}, \mathcal{B}]$  and identically equal to 1 on  $[-\mathcal{B}^{-1}, \mathcal{B}^{-1}]$  such that*

$$b_{j,\ell} = \mathbf{b}(\mathcal{B}^{-j}\ell)$$

where  $\mathbf{b}(\cdot) = \sqrt{\mathbf{a}(\cdot/\mathcal{B}) - \mathbf{a}(\cdot)}$ .

For such window functions,  $\tilde{\ell}_j^{max} = \mathcal{B}^{j+1}$ . These spectral windows are not as general as those used by Faÿ *et al.* [78]. Indeed, it has been shown by Guillaux *et al.* [25] that one can take advantage of the relaxation of the  $\mathcal{B}$ -adic scheme originally proposed in the definition of needlets to optimize their non-asymptotic localization properties. In the following, since we are concerned with asymptotic properties, we will make use of the  $\mathcal{B}$ -adic structure of Assumptions 7.1 and 7.2, so that the spatial localization property of the needlet takes the convenient form of the next proposition.

**Proposition 7.3** ([33]). *There exists a constant  $c > 0$  which depends only on the function  $\mathbf{b}$  such that for all  $j \in \mathbb{N}$ ,  $k \in \{1, \dots, N_j\}$  and  $\xi \in \mathbb{S}^2$ ,*

$$|\psi_{j,k}(\xi)| \leq \frac{c\mathcal{B}^j}{(1 + \mathcal{B}^j d(\xi, \xi_{j,k}))^M} \quad .$$

The stochastic counterpart of this analytical result is that the needlet coefficients of a stationary field are asymptotically uncorrelated as  $j \rightarrow \infty$ , except for points at a distance of order  $\mathcal{B}^{-j}$  or less. For this purpose and throughout this article, we make on the power spectrum of  $X$  the same following regularity assumption as in [151, 171, 170] (and also, in substance, in [175]).

**Assumption 7.2.** *There exist  $\alpha > 2$  and a sequence of functions  $\mathbf{g}_j : [\mathcal{B}^{-1}, \mathcal{B}] \rightarrow \mathbb{R}$ ,  $j \in \mathbb{N}$ , such that*

$$C_\ell = \ell^{-\alpha} \mathbf{g}_j(\mathcal{B}^{-j}\ell)$$

*for every  $\ell \in [\mathcal{B}^{j-1}, \mathcal{B}^{j+1}]$ . Moreover, there exist positive numbers  $c_0, \dots, c_M$  such that for all  $j \in \mathbb{N}$ ,  $c_0^{-1} \leq \mathbf{g}_j \leq c_0$  and for all  $r \leq M$ ,  $\sup_{\mathcal{B}^{-1} \leq u \leq \mathcal{B}} \left| \frac{d^r}{du^r} \mathbf{g}_j(u) \right| \leq c_r$ .*

**Proposition 7.4** ([151]). *Let  $X$  be stationary with a power spectrum satisfying Assumption 7.2 and  $\eta_{j,k}$  its needlet coefficients. Then there exist a constant  $c > 0$  such that for all  $j \in \mathbb{N}$  and  $k, k' \in \{1, \dots, N_j\}$*

$$|\text{Cor}[\eta_{j,k}, \eta_{j,k'}]| \leq \frac{c}{(1 + \mathcal{B}^j d(\xi_{j,k}, \xi_{j,k'}))^M} \quad .$$

**Remark.** *A generalization of this  $\mathcal{B}$ -adic framework, in a different direction to the one of Guillaux et al. [25] can be found in Geller & Mayeli [21, 175] and Lan & Marinucci [170]. The authors do not suppose that the function  $\mathbf{a}$  (or  $\mathbf{b}$ ) is compactly supported and obtain localization and asymptotic uncorrelation results similar to Propositions 7.3 and 7.4.*

### 7.3 Estimation results

In this Section, we present a new procedure for the estimation of the angular power spectrum of  $X$  in the model of Eq. (7.2) based on the needlet coefficients of  $Y$ . The properties of needlets described in Section 7.2.4 allow to take into account the local signal-to-noise ratio in the estimation of the (however) global spectrum. This spatial accuracy is at the cost of a lower frequential precision : not every value of  $C_\ell$  will be estimated, but only the mean values in the bands defined by the windows  $b_{j,\ell}$ .

### 7.3.1 Parameters and estimators

We shall provide below estimators for the parameters

$$C^{(j)} := (4\pi)^{-1} \sum_{\ell \geq 0} (2\ell + 1) (b_{j,\ell})^2 C_\ell$$

which are smoothed versions of the power spectrum of  $X$  at successive scales. A comparison of a Riemannian sum to an integral shows that, under Assumption 7.2,  $c_1 \mathcal{B}^{j(2-\alpha)} \leq C^{(j)} \leq c_2 \mathcal{B}^{j(2-\alpha)}$  for some  $c_1, c_2 > 0$ .

Following [149, 170, 171], we consider the “consistency” of a sequence of estimators in the high-frequency asymptotics, which implies (contrary to the usual notion of consistency) a sequence of parameters to estimate. This is made more precise in the definition below.

**Définition.** Given a sequence  $\theta_j$  of nonzero parameters, a sequence  $\hat{\theta}_j$  of estimators is said consistent if  $\mathcal{E}(\theta_j, \hat{\theta}_j) \xrightarrow{j \rightarrow \infty} 0$ , where

$$\mathcal{E}(\theta_j, \hat{\theta}_j) := \frac{\mathbf{E}(\hat{\theta}_j - \theta_j)^2}{\theta_j^2}.$$

The Needlet Spectral Estimators (NSE) are defined by

$$\widehat{C}^{(j)} := \sum_{k=1}^{N_j} w_{j,k} \left[ (\gamma_{j,k})^2 - (n_{j,k})^2 \right] \quad (7.5)$$

with some weights such that  $\sum_{k=1}^{N_j} w_{j,k} = 1$  and where

$$n_{j,k} := \left( \frac{1}{\lambda_{j,k}} \sum_{p=1}^{N_j} (\lambda_{j,p} \sigma_{j,p} \psi_{j,k}(\xi_p))^2 \right)^{1/2} \quad (7.6)$$

is the standard deviation of the needlet coefficients of the masked noise (it follows from Proposition 7.2 with  $k = k'$  and  $W_{j,k} \sigma_{j,k}$  in place of  $\sigma_{j,k}$ ). The  $\gamma_{j,k}$ 's are the observed needlet coefficients of  $Y$  and  $n_{j,k}$  are defined in Eq. (7.6). The weights should ideally be designed according to the local noise level and the local effect of the mask. For this purpose, given a sequence  $(t_j)_{j \in \mathbb{N}}$  in  $\mathbb{R}$ , let us define the set of *kept coefficients*

$$\mathcal{K}_j := \left\{ k \in \{1, \dots, N_j\} : \left\{ \sum_{p=1}^{N_j} \lambda_{j,p} (1 - W_{j,p})^2 (\psi_{j,k}(\xi_{j,p}))^2 \right\}^{1/2} \leq t_j \right\}.$$

As the second-order characteristics of the noise are known, its effect on the square of the needlet coefficients is completely subtracted in (7.5). Thus, the bias of the NSE is caused only by the mask.

**Proposition 7.5.** For any choice of weights such that  $w_{j,k} = 0$  for  $k \notin \mathcal{K}_j$ , there exists a constant  $c$  such that

$$\frac{|\mathbf{E} \widehat{C}^{(j)} - C^{(j)}|}{C^{(j)}} \leq c \mathcal{B}^{\alpha_j} t_j.$$

**Remark 7.1.** In particular, if  $W_j \equiv 1$  (no mask), then for any choice of weights,  $\widehat{C}^{(j)}$  is unbiased.

This result is completed by the following one about the variance of  $\widehat{C}$  when there is no noise.

**Proposition 7.6.** *Suppose that  $\sigma_{j,k} = 0$  for all  $k \in \{1, \dots, N_j\}$  and that  $w_{j,k} = 0$  for  $k \notin \mathcal{K}_j$ . Then there exists  $c > 0$  such that*

$$\frac{\mathbf{Var}(\widehat{C}^{(j)})}{C^2} \leq c \left\{ \sum_{k,k'=1}^{N_j} \frac{|w_{j,k}w_{j,k'}|}{(1 + \mathcal{B}^j d(\xi_{j,k}, \xi_{j,k'}))^{2M}} + \frac{N_j t_j}{C^{(j)}} + \left( \frac{N_j t_j}{C^{(j)}} \right)^2 \right\}$$

If the needlets coefficients at a scale  $j$  were independent, centered Gaussian variables with unknown variance  $C^{(j)}$  and observed with independent centered heteroscedastic Gaussian additive errors of variance  $(n_{j,k})^2$ , then the Maximum Likelihood of  $C^{(j)}$  would be reached for the value  $\widehat{C}^{(j)}$  given by (7.5) with weights  $w_{j,k} \propto (C^{(j)} + (n_{j,k})^2)^{-2}$  (see e.g. Faÿ *et al.* [78]). In reality, the uncorrelation of the needlets coefficients holds only approximately in the sens of Proposition 7.4. However, if some previous, possibly rough, estimates for  $C^{(j)}$  are available, say  $\overline{C}^{(j)}$ , the following weights are still meaningful.

$$w_{j,k} = \frac{1}{S} \left( \overline{C}^{(j)} + (n_{j,k})^2 \right)^{-2} \mathbf{I}(k \in \mathcal{K}_j) \quad (7.7)$$

where  $S := \sum_{k=1}^{N_j} (\overline{C}^{(j)} + (n_{j,k})^2)^{-2} \mathbf{I}(k \in \mathcal{K}_j)$ .

### 7.3.2 Consistency of the estimators

Suppose that Assumptions 7.1 and 7.2 are satisfied. Theorems 7.1 and 7.2 below provide conditions for the consistency of the needlet spectral estimators (7.5). Theorem 7.1 deals with generic weights whereas in Theorem 7.2 we consider weights of the form (7.7).

**Theorem 7.1.** *Assume that*

i) *the weights  $w_{j,k}$  are such that  $w_{j,k} = 0$  for  $k \notin \mathcal{K}_j$  and that moreover*

$$\sum_{k,k'=1}^{N_j} \frac{|w_{j,k}w_{j,k'}|}{(1 + \mathcal{B}^j d(\xi_{j,k}, \xi_{j,k'}))^M} \xrightarrow{j \rightarrow \infty} 0$$

and

$$\sup_{1 \leq p \leq N_j} \sum_{k,k'=1}^{N_j} \frac{|w_{j,k}w_{j,k'}|}{(1 + \mathcal{B}^j d(\xi_{j,p}, \xi_{j,k}))^{2M} (1 + \mathcal{B}^j d(\xi_{j,p}, \xi_{j,k'}))^{2M}} \xrightarrow{j \rightarrow \infty} o(\mathcal{B}^{-4j}) ;$$

ii) *the noise verifies*

$$\sum_{p=1}^{N_j} \lambda_{j,p} (W_{j,p} \sigma_{j,p})^4 \xrightarrow{j \rightarrow \infty} o(\mathcal{B}^{(4-\alpha)j}) ;$$

iii) *the thresholds  $t_j$  defining the set of kept coefficients are chosen such that*

$$t_j \xrightarrow{j \rightarrow \infty} o(\mathcal{B}^{-\alpha j}) .$$

*Then the sequence  $\widehat{C}^{(j)}$  is consistent.*

**Remark 7.2.** Condition i) prohibits weights which are too much concentrated on few points  $k$ . For instance, thanks to Lemma 7.1, Condition i) is trivially satisfied if

$$\max_{1 \leq k \leq N_j} |w_{j,k}| \underset{j \rightarrow \infty}{=} O(N_j^{-1}).$$

**Theorem 7.2.** Suppose that the weights follow Eq. (7.7) and that

i) there exists a constant  $\kappa_1 > 1$  and a sequence  $r_j$  in  $\mathbb{R}_+$  such that

$$\frac{(1 + r_j)^2}{\text{Card} \{k \in \mathcal{K}_j : (n_{j,k})^2 \leq r_j C^{(j)}\}} \underset{j \rightarrow \infty}{=} O(\mathcal{B}^{-\kappa_1 j});$$

ii) there exists a constant  $\kappa_2 \in \mathbb{R}$  such that

$$\sum_{p=1}^N \lambda_p (W_{j,p} \sigma_{j,p})^2 \underset{j \rightarrow \infty}{=} O(\mathcal{B}^{\kappa_2 j})$$

iii) the thresholds  $t_j$  defining the set of kept coefficients are such that

$$t_j \underset{j \rightarrow \infty}{=} o(\mathcal{B}^{-\alpha j});$$

iv) there exists a constant  $c > 0$  such that for all  $j \in \mathbb{N}$

$$c^{-1} C^{(j)} \leq \bar{C}^{(j)} \leq c C^{(j)}.$$

Then, for  $M > \max \left\{ \frac{\kappa_2 + \alpha}{\kappa_1 - 1}; \frac{5 + \alpha - \kappa_1 + \kappa_2}{\kappa_1 - 1} \right\}$ , the sequence  $\widehat{C}^{(j)}$  is consistent.

The hypotheses of Theorem 7.2 can be interpreted in the following way : Condition i) describes a trade-off between an arbitrary upper bound on the noise variance and the number of observed points under this bound. The noise level can be as high as wanted (in the – very weak – limit of Condition ii)) in some regions, provided that it is under control at least in some other regions. An example is given below. Conditions ii) and iii) are not restrictive, neither is the condition on the number  $M$  of derivatives of  $\mathbf{a}$ , since one is free to chose a infinitely differentiable  $\mathbf{a}$ . On the contrary, Condition iv) assumes a previous partial knowledge on the parameters  $C^{(j)}$  to estimate, which is equivalent to the knowledge of the parameter  $\alpha$ .

**Example.** Suppose that there exists a  $\delta > 0$  such that for each scale  $j$ , data with a noise variance less than  $\sqrt{N_j}$  are available at least within a spherical disc of radius  $\delta$ . More precisely, there exist  $\xi_j^* \in \mathbb{S}^2$

and  $r_j = O\left(\left(\sqrt{N_j}\right)^\beta\right)$ ,  $\beta < 1$  such that for all  $k$  with  $d(\xi_{j,k}, \xi_j^*) \leq \delta$

- $W(\xi_{j,k}) = 1$
- and  $\sigma_{j,k} \leq r_j C^{(j)}$ .

Then condition i) is satisfied.

## 7.4 Simulations

In this section, we investigate the non-asymptotic numerical performances of the NSE, in a relatively realistic setting. This is for illustration only, as there are many free parameters in the models (mask, noise level) that are only asymptotically constrained by the previous theory.



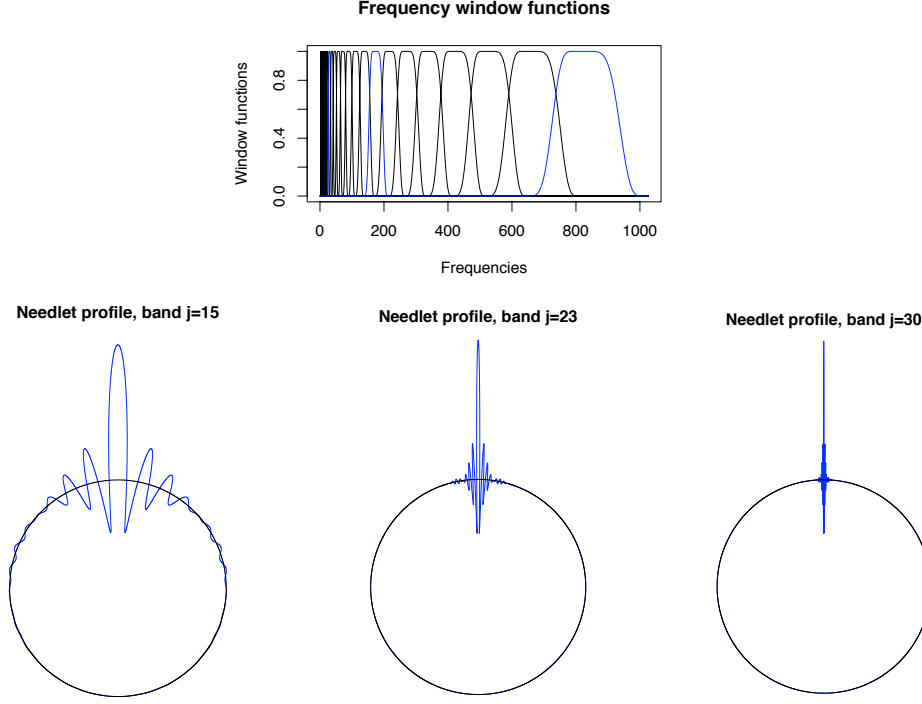


FIG. 7.1 – Needlets in frequency and spatial spaces. Top : window functions  $b_{j,\ell}$  as functions of frequency  $\ell$  (in blue, the three ones corresponding to profiles plotted below). Bottom : Polar plots of the function  $\theta \in [0, 2\pi[ \mapsto \sum_{\ell \geq 0} b_{j,\ell} L_\ell(\cos \theta)$  for  $j = 15, 23$  and  $30$ , to illustrate their localization property.

#### 7.4.1 Settings

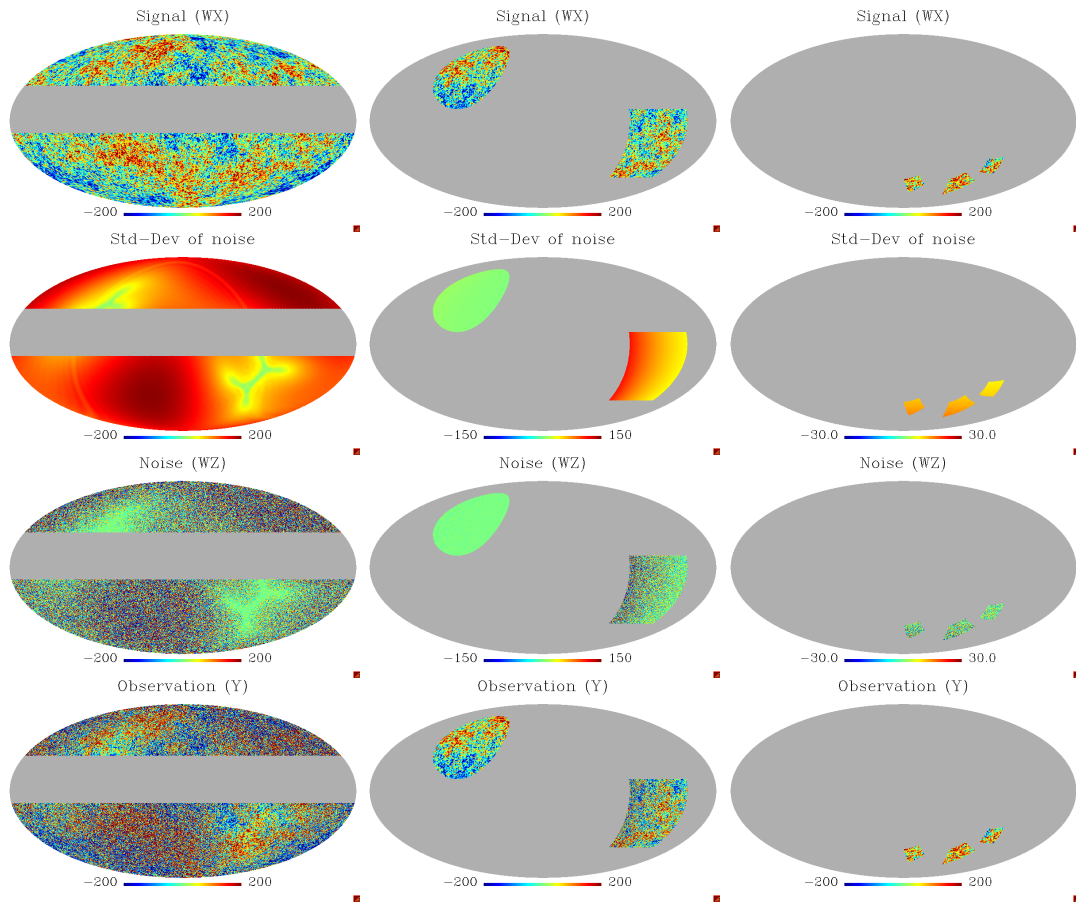
The constant determining the width of the bands is set to  $\mathcal{B} = 1.25$ . The function  $\mathbf{a}$  of Assumption 7.1 is designed on  $[\mathcal{B}^{-1}, \mathcal{B}]$  by a polynomial of degree 19 connecting 1 to 0 with  $M = 9$  continuous derivatives, *i.e.*  $\mathbf{a}(\mathcal{B}^{-1}) = 1$ ,  $\mathbf{a}(\mathcal{B}) = 0$  and  $\mathbf{a}^{(r)}(\mathcal{B}^{\pm 1}) = 0$ ,  $1 \leq r \leq M$ . Frequency window functions  $b_{j,\ell}$  and some of the associated needlets are displayed in Fig. 7.1.

In the analysis, the first 10 scales ( $0 \leq j \leq 9$ ) are not considered since the support of the corresponding window functions is of width less or equal than 2. These scales correspond to frequencies  $\ell \leq 8$ . The range of studied scales is  $10 \leq j \leq 30$ .

We use the HEALPix pixelization and associated direct and inverse spherical harmonic transform [84]. Needlet coefficients are computed from three experiments providing maps at respective HEALPix resolutions  $\text{nside} = 128, 256$  and  $512$  (*i.e.* the number of cubature points  $N$  is 196.608, 786.432 and 3.145.728).

The simulated data are illustrated in Fig. 7.2. A first “experiment” (hereafter,  $A$ ), *i.e.* a mask and a noise level map, is used for scales  $10 \leq j \leq 23$ . The noise level is reminiscent to the expected level from the forthcoming Planck experiment of the European Spatial Agency<sup>1</sup>. Two other experiments (hereafter,  $B$  and  $C$ ), with synthetic masks and noise-level maps, are used respectively for scales  $23 < j \leq 26$  and  $26 < j \leq 30$ .

<sup>1</sup>see [www.rssd.esa.int/Planck/](http://www.rssd.esa.int/Planck/)



F . 7.2 – Simulated data. First line :  $WX$ . Second line :  $W\sigma$ . Third line :  $WZ$ . Fourth line :  $Y$ . The columns correspond to the settings for the three experiments : A (first column) used for  $10 \leq j \leq 23$  ; B (second column) used for  $23 < j \leq 26$  ; C (third column) used for  $26 < j \leq 30$ .

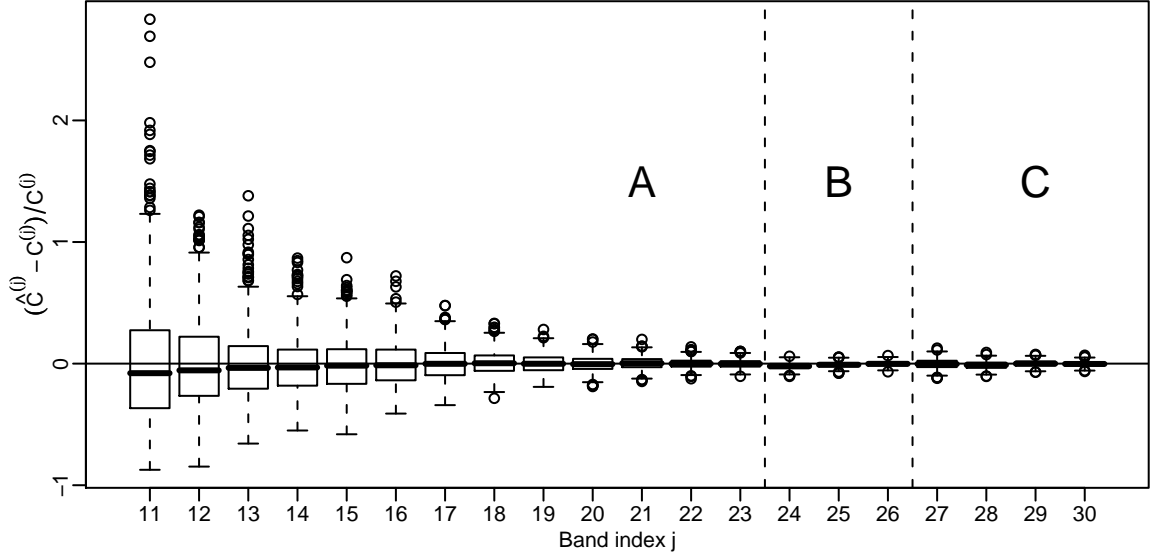


FIG. 7.3 – Box-and-whisker plots of the NSE estimator  $\widehat{C}^{(j)}$ . Regions A,B, and C delimits the range of the three experiments, from the coarsest (and wide coverage) to the finest (and small coverage) one.

#### 7.4.2 Results

The distribution on the NSE estimators for first the scales  $10 \leq j \leq 30$  is estimated by 800 Monte Carlo replicates in the setting of previous section. Figure 7.3 illustrates the convergence of the sequence  $\widehat{C}^{(j)}$ . The approximate normal distribution of the estimator is highlighted by the quantile-to-quantile plots of Figure 7.4 and the  $p$ -values from Anderson-Darling goodness-of-fit test.

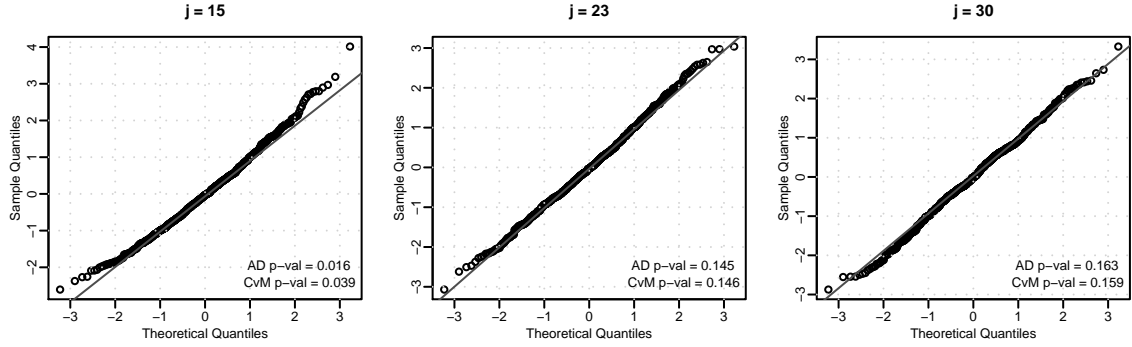
### 7.5 Conclusion

In this work, we have proved the mean-square consistency of a spherical spectral estimator based on the needlet analysis of a stationary field with missing data and non-homogeneous additive noise. Our model is relatively realistic and could be enriched to fit the requirement of particular applications. A possible development of this work is the study of the rate of convergence, limiting law, efficiency, etc. of those estimators.

Although this has not been stated here for brevity, it must be noticed that this method allows easy adaptation to a context with longitudinal data, *i.e.* of the form

$$Y_{j,k,e} = W_{j,k,e} \left[ X(\xi_{j,k}) + Z_{j,k,e} \right], \quad j \in \mathbb{N}, \quad k \in \{1, \dots, N_j\}, \quad e \in \{1, \dots, E_j\}.$$

for some sequence of integer  $\{E_j\}_{j \in \mathbb{N}}$ . This generalization of model (7.2) is of high importance for practical applications, when for each scale  $j$  multiple independent observations (indexed by  $e$ ) of the signal of interest are available, such those provided by the various CMB experiments (see Faÿ *et al.* [78] for details).



F . 7.4 – Distribution of the standardized NSE estimator  $\widehat{C}^{(j)}$  for bands 15, 23 and 30. The  $p$ -values are obtained from the Anderson-Darling and Cramér-von Mises tests of Gaussianity. At low frequencies, the estimator is significantly skewed.

## 7.6 Proofs

In this section, for sake of notational simplicity,

- the sub- and superscripts  $j$  are omitted
- for quantities  $a(j)$  and  $b(j)$  depending on  $j$ ,  $a \lesssim b$  means :  $\exists c > 0, \forall j \in \mathbb{N}, a(j) \leq cb(j)$
- $a \asymp b$  means :  $a \lesssim b$  and  $b \lesssim a$ .

The following fact will be used extensively in the proofs. Let  $(\xi_p, \lambda_p)_{1 \leq p \leq N}$  be a pixelization of order  $4\mathcal{B}^{j+1}$

$$\begin{aligned}
 \sum_{p=1}^N \lambda_p \psi_k^2(\xi_p) &= \int_{\mathbb{S}^2} \psi_k^2(\xi) d\xi = \lambda_k^{-1} \sum_{\ell, \ell' \geq 0} \mathbf{b}(\mathcal{B}^{-j}\ell) \mathbf{b}(\mathcal{B}^{-j}\ell') \int_{\mathbb{S}^2} L_\ell(\xi_k \cdot \xi) L_{\ell'}(\xi_k \cdot \xi) d\xi \\
 &= \lambda_k^{-1} \sum_{\ell \geq 0} \mathbf{b}^2(\mathcal{B}^{-j}\ell) \frac{2\ell+1}{4\pi} \asymp \lambda_k^{-1} \mathcal{B}^j \sum_{\ell \geq 0} \mathbf{b}^2(\mathcal{B}^{-j}\ell) \mathcal{B}^{-j}\ell \asymp 1
 \end{aligned} \tag{7.8}$$

as  $\int_{\mathbb{R}_+} \mathbf{b}(x) x dx > 0$  and  $\lambda_k^{-1} \mathcal{B}^j \asymp \mathcal{B}^{-j}$  using (7.1b)-(7.1c).

### 7.6.1 Proof of Propositions 7.1 and 7.2

The four following lines, which can be found in [151, p. 9], are recalled for the reader's convenience.

$$\begin{aligned}
 \mathbf{E} [\eta_k \eta_{k'}] &= (\lambda_k \lambda_{k'})^{-1/2} \mathbf{E} \iint_{\mathbb{S}^2} X(\xi) X(\xi') \psi_k(\xi) \psi_{k'}(\xi') d\xi d\xi' \\
 &= (\lambda_k \lambda_{k'})^{-1/2} \mathbf{E} \iint_{\mathbb{S}^2} \sum_{\ell, \ell' \geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} a_{\ell, m} a_{\ell', m'} \mathcal{Y}_{\ell, m}(\xi) \mathcal{Y}_{\ell', m'}(\xi') \psi_k(\xi) \psi_{k'}(\xi') d\xi d\xi' \\
 &= (\lambda_k \lambda_{k'})^{-1/2} \sum_{\ell \geq 0} C_\ell \sum_{m=-\ell}^{\ell} \int_{\mathbb{S}^2} \mathcal{Y}_{\ell, m}(\xi) \psi_k(\xi) d\xi \int_{\mathbb{S}^2} \mathcal{Y}_{\ell, m}(\xi) \psi_{k'}(\xi) d\xi \\
 &= \sum_{\ell \geq 0} C_\ell b_\ell^2 \sum_{m=-\ell}^{\ell} \mathcal{Y}_{\ell, m}(\xi_k) \mathcal{Y}_{\ell, m}(\xi_{k'})
 \end{aligned}$$

which proves Proposition 7.1. On the other hand  $\zeta_k = \sum_{\ell \geq 0} \sum_{m=-\ell}^{\ell} b_{\ell} z_{\ell,m} \mathcal{Y}_{\ell,m}(\xi_k)$  where

$$z_{\ell,m} := \sum_{p=1}^N \lambda_p Z_p \mathcal{Y}_{\ell,m}(\xi_p)$$

Thus

$$\begin{aligned} \mathbf{Cov}[\zeta_k \zeta_{k'}] &= \sum_{\ell, \ell' \geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} b_{\ell} b_{\ell'} \mathbf{Cov}[z_{\ell,m}, z_{\ell',m'}] \mathcal{Y}_{\ell,m}(\xi_k) \mathcal{Y}_{\ell',m'}(\xi_{k'}) \\ &= \sum_{\ell, \ell' \geq 0} \sum_{m=-\ell}^{\ell} \sum_{m'=-\ell'}^{\ell'} b_{\ell} b_{\ell'} \left( \sum_{p=1}^N \lambda_p^2 \sigma_p^2 \mathcal{Y}_{\ell,m}(\xi_p) \mathcal{Y}_{\ell',m'}(\xi_p) \right) \mathcal{Y}_{\ell,m}(\xi_k) \mathcal{Y}_{\ell',m'}(\xi_{k'}) \\ &= \frac{1}{\sqrt{\lambda_k \lambda_{k'}}} \sum_{p=1}^N \lambda_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \end{aligned}$$

## 7.6.2 Proof of Proposition 7.5

$$\begin{aligned} \frac{|\mathbf{E} \widehat{C} - C|}{C} &= \frac{1}{C} \left| \sum_{k \in \mathcal{K}} w_k \{ \mathbf{Var}(\gamma_k) - n_k^2 - C \} \right| \\ &= \frac{1}{C} \left| \sum_{k \in \mathcal{K}} w_k \{ \mathbf{Var}(\eta_k) - C \} \right| \\ &\leq \frac{1}{C} \sum_{k \in \mathcal{K}} w_k \lambda_k^{-1} \left| \mathbf{E} \left( \sum_{p=1}^N \lambda_p W_p X(\xi_p) \psi_k(\xi_p) \right)^2 - \mathbf{E} \left( \sum_{p=1}^N \lambda_p X(\xi_p) \psi_k(\xi_p) \right)^2 \right| \\ &= \frac{1}{C} \sum_{k \in \mathcal{K}} w_k \lambda_k^{-1} \left| \mathbf{E} \left( \sum_{p=1}^N \lambda_p (1 - W_p) X(\xi_p) \psi_k(\xi_p) \sum_{p=1}^N \lambda_p (1 + W_p) X(\xi_p) \psi_k(\xi_p) \right) \right| \\ &\leq \frac{2}{C} \sum_{k \in \mathcal{K}} w_k \lambda_k^{-1} \left\{ \mathbf{E} \left( \sum_{p=1}^N \lambda_p (1 - W_p) X(\xi_p) \psi_k(\xi_p) \right)^2 \mathbf{E} \left( \sum_{p=1}^N \lambda_p X(\xi_p) \psi_k(\xi_p) \right)^2 \right\}^{1/2} \\ &\leq \frac{8\pi \mathbf{Var}(X)}{C} \sum_{k \in \mathcal{K}} w_k \lambda_k^{-1} \left\{ \sum_{p=1}^N \lambda_p (1 - W_p)^2 \psi_k^2(\xi_p) \sum_{p=1}^N \lambda_p \psi_k^2(\xi_p) \right\}^{1/2} \\ &\lesssim \frac{N}{C} t \sum_{k \in \mathcal{K}} w_k \\ &\asymp \mathcal{B}^{\alpha j_t} . \end{aligned}$$

In particular, if  $W \equiv 1$ , one can take  $t = 0$  in the definition of  $\mathcal{K}$  and the above result shows that  $\widehat{C}$  is unbiased.

### 7.6.3 Proof of Proposition 7.6

Define the unobserved coefficients  $\tilde{\eta}_k := \lambda_k^{-1/2} \langle \psi_k, X \rangle$  .

$$\begin{aligned}
\frac{\mathbf{Var}(\widehat{C})}{C^2} &= \frac{1}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}[\eta_k^2, \eta_{k'}^2] = \frac{2}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}^2[\eta_k, \eta_{k'}] \\
&= \frac{2}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov}[\tilde{\eta}_k, \tilde{\eta}_{k'}] - \mathbf{Cov}[\tilde{\eta}_k, \tilde{\eta}_{k'} - \eta_{k'}] \\
&\quad - \mathbf{Cov}[\tilde{\eta}_k - \eta_k, \tilde{\eta}_{k'}] + \mathbf{Cov}[\tilde{\eta}_k - \eta_k, \tilde{\eta}_{k'} - \eta_{k'}] \}^2 \\
&\asymp \frac{1}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov}^2[\tilde{\eta}_k, \tilde{\eta}_{k'}] + \mathbf{Cov}^2[\tilde{\eta}_k, \tilde{\eta}_{k'} - \eta_{k'}] + \mathbf{Cov}^2[\tilde{\eta}_k - \eta_k, \tilde{\eta}_{k'} - \eta_{k'}] \} \quad (7.9)
\end{aligned}$$

For the first term in the r.h.s. of Eq. (7.9), we use the following bound, proved in [151].

$$\sum_{\ell \geq 0} b_\ell^2 C_\ell L_\ell(\xi_k \cdot \xi_{k'}) \lesssim \frac{\mathcal{B}^{j(2-\alpha)}}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^M} \quad (7.10)$$

Thus, with Lemma 7.1 and using that  $C \asymp \mathcal{B}^{j(2-\alpha)}$  and  $N \asymp \mathcal{B}^{2j}$

$$\begin{aligned}
\sum_{k,k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}^2[\tilde{\eta}_k, \tilde{\eta}_{k'}] &= \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \left\{ \sum_{\ell \geq 0} b_\ell^2 C_\ell L_\ell(\xi_k \cdot \xi_{k'}) \right\}^2 \\
&\lesssim \mathcal{B}^{2j(2-\alpha)} \sum_{k,k' \in \mathcal{K}} \frac{|w_k w_{k'}|}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} \\
&\lesssim C^2 \sum_{k,k' \in \mathcal{K}} \frac{|w_k w_{k'}|}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} \quad .
\end{aligned}$$

For the other terms of the decomposition (7.9), one can write

$$\begin{aligned}
\mathbf{Var}(\tilde{\eta}_k - \eta_k) &= \lambda_k^{-1} \mathbf{E} \left( \sum_{p=1}^N \lambda_p (1 - W_p) X(\xi_p) \psi_k(\xi_p) \right)^2 \\
&\leq \lambda_k^{-1} \mathbf{E} \left( \sum_{p=1}^N \lambda_p X^2(\xi_p) \sum_{p=1}^N \lambda_p (1 - W_p) \psi_k^2(\xi_p) \right) \\
&= \lambda_k^{-1} 4\pi \mathbf{Var}(X) \sum_{p=1}^N \lambda_p (1 - W_p) \psi_k^2(\xi_p) \\
&\lesssim Nt \quad (7.11)
\end{aligned}$$

thus

$$\begin{aligned}
\sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov}[\tilde{\eta}_k - \eta_k, \tilde{\eta}_{k'}] \}^2 &\leq \sum_{k \in \mathcal{K}} w_k \mathbf{Var}(\tilde{\eta}_k - \eta_k) \sum_{k' \in \mathcal{K}} w_{k'} \mathbf{Var}(\tilde{\eta}_{k'}) \\
&\lesssim CNt
\end{aligned}$$

and

$$\begin{aligned} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov} [\tilde{\eta}_k - \eta_k, \tilde{\eta}_{k'} - \eta_{k'}] \}^2 &\leq \left\{ \sum_{k \in \mathcal{K}} w_k \mathbf{Var} (\tilde{\eta}_k - \eta_k) \right\}^2 \\ &\lesssim N^2(t)^2 . \end{aligned}$$

Putting these terms together establishes Proposition 7.6.

#### 7.6.4 Proof of Theorems 7.1 and 7.2

We have

$$\mathcal{E}(\widehat{C}) = \frac{|\mathbf{E} \widehat{C} - C|^2}{C^2} + \frac{\mathbf{Var}(\widehat{C})}{C^2} .$$

By Proposition 7.5, the bias term is bounded by

$$\frac{|\mathbf{E} \widehat{C} - C|}{C} \lesssim \mathcal{B}^{\alpha} j_t$$

and converges to 0 thanks to Condition iii).

On the other hand

$$\begin{aligned} \frac{\mathbf{Var}(\widehat{C})}{C^2} &= \frac{1}{C^2} \sum_{k,k'=1}^N w_k w_{k'} \mathbf{Cov} [\gamma_k^2, \gamma_{k'}^2] = \frac{2}{C^2} \sum_{k,k'=1}^N w_k w_{k'} \mathbf{Cov}^2 [\gamma_k, \gamma_{k'}] \\ &= \frac{2}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov} [\eta_k, \eta_{k'}] + \mathbf{Cov} [\zeta_k, \zeta_{k'}] \}^2 \\ &\asymp \frac{1}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \{ \mathbf{Cov}^2 [\eta_k, \eta_{k'}] + \mathbf{Cov}^2 [\zeta_k, \zeta_{k'}] \} . \end{aligned} \quad (7.12)$$

From Proposition 7.6

$$\frac{1}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}^2 [\eta_k, \eta_{k'}] \lesssim \sum_{k,k'=1}^N \frac{|w_k w_{k'}|}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} + \frac{Nt}{C} + \left( \frac{Nt}{C} \right)^2$$

where  $\frac{Nt}{C} \lesssim \mathcal{B}^{\alpha} j_t \xrightarrow{j \rightarrow \infty} 0$  with Condition iii). To establish Theorems 7.1 and 7.2 it suffices now to prove that in both cases

$$\sum_{k,k'=1}^N \frac{|w_k w_{k'}|}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} \xrightarrow{j \rightarrow \infty} 0 \quad (7.13)$$

and

$$\frac{1}{C^2} \sum_{k,k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}^2 [\zeta_k, \zeta_{k'}] \xrightarrow{j \rightarrow \infty} 0 . \quad (7.14)$$

Let us consider separately the cases of Theorems 7.1 and 7.2.

### 7.6.5 End of proof of Theorem 7.1

The convergence of Eq. (7.13) is nothing else but Condition i). To prove (7.14), use the Property (7.1c) and the Cauchy-Schwartz inequality and write

$$\begin{aligned}
\sum_{k,k' \in \mathcal{K}} |w_k w_{k'}| \mathbf{Cov}^2[\zeta_k, \zeta_{k'}] &= \sum_{k,k' \in \mathcal{K}} \frac{|w_k w_{k'}|}{\lambda_k \lambda_{k'}} \left\{ \sum_{p=1}^N \lambda_p^2 W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\
&\asymp \sum_{k,k' \in \mathcal{K}} |w_k w_{k'}| \left\{ \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\
&\lesssim \sum_{k,k' \in \mathcal{K}} |w_k w_{k'}| \left\{ \sum_{p=1}^N \lambda_p W_p^4 \sigma_p^4 \right\} \left\{ \sum_{p=1}^N \lambda_p |\psi_k(\xi_p) \psi_{k'}(\xi_p)|^2 \right\} \\
&\lesssim \left\{ \sum_{p=1}^N \lambda_p W_p^4 \sigma_p^4 \right\} \sum_{k,k' \in \mathcal{K}} |w_k w_{k'}| \\
&\quad \times \mathcal{B}^{2j} \sum_{p=1}^N \frac{1}{(1 + \mathcal{B}^j d(\xi_p, \xi_k))^{2M} (1 + \mathcal{B}^j d(\xi_p, \xi_{k'}))^{2M}} \\
&\stackrel{j \rightarrow \infty}{=} o(C)
\end{aligned}$$

with Conditions i) and ii) and Proposition 7.3.

### 7.6.6 End of proof of Theorem 7.2

The two following remarks will prove useful. First, from Condition iv),  $(c^{-1} \wedge 1)(C + n_k^2) \leq (\bar{C} + n_k^2) \leq (c \vee 1)(C + n_k^2)$  so that

$$w_k \asymp \frac{1}{S} (C + n_k^2)^{-2} \mathbf{I}(k \in \mathcal{K}) \quad . \quad (7.15)$$

Second, using again Condition iv)

$$S = \sum_{k \in \mathcal{K}} (\bar{C} + n_k^2)^{-2} \geq \sum_{k \in V} (\bar{C} + n_k^2)^{-2} \geq \frac{\text{Card } V}{(1+r)^2 \bar{C}^2} \gtrsim \frac{\text{Card } V}{(1+r)^2 C^2} \quad (7.16)$$

where

$$V := \{k \in \mathcal{K} : n_k^2 \leq rC\} \quad . \quad (7.17)$$

The convergence of Eq. (7.13) is established thanks to Conditions i) and iii) and Lemma 7.1

$$\begin{aligned}
\sum_{k,k'=1}^N \frac{|w_k w_{k'}|}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} &\asymp \frac{1}{S^2} \sum_{k,k' \in \mathcal{K}} \frac{(C + n_k^2)^{-2} (C + n_{k'}^2)^{-2}}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} \\
&\leq \frac{C^{-4}}{S^2} \sum_{k,k' \in \mathcal{K}} \frac{1}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^{2M}} \\
&\lesssim \frac{(1+r)^4}{(\text{Card } V)^2} (\text{Card } \mathcal{K}) \xrightarrow{j \rightarrow \infty} 0
\end{aligned}$$



For the convergence of Eq. (7.14), note that  $\mathbf{Cov}[\zeta_k, \zeta_{k'}] \asymp \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p)$  where the implicit constant does not depend on  $k, k'$  (Proposition 7.2 and property (7.1c) of the pixelization). For any constant  $\kappa \in (0, 1)$ , the value of which will be fixed later, we have

$$\begin{aligned} \frac{1}{C^2} \sum_{k, k' \in \mathcal{K}} w_k w_{k'} \mathbf{Cov}^2[\zeta_k, \zeta_{k'}] &\asymp \frac{1}{C^2 S^2} \sum_{k, k' \in \mathcal{K}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \\ &\quad \times \left\{ \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\ &\lesssim \frac{1}{C^2 S^2} (A_1 + A_2 + A_3 + A_4) \end{aligned}$$

with

$$\begin{aligned} A_1 &= \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) \leq \mathcal{B}^{-\kappa j}}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\ A_2 &= \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ \sum_{p \in D(k, k')} \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\ &\quad \text{where } D(k, k') := \{p : d(\xi_k, \xi_p) > \frac{1}{2} \mathcal{B}^{-\kappa j} \text{ and } d(\xi_{k'}, \xi_p) > \frac{1}{2} \mathcal{B}^{-\kappa j}\} \\ A_3 &= \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ n_k^2 \sum_{p: d(\xi_k, \xi_p) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}} \lambda_p \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\ A_4 &= \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ \sum_{p: d(\xi_k, \xi_p) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}} \lambda_p |W_p^2 \sigma_p^2 - n_k^2| \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2. \end{aligned}$$

**Remark.** As noticed in Section 7.3.1, the weights proposed here are linked with the asymptotic uncorrelation of the needlet coefficients.  $A_1$  includes the near-diagonal terms of the noise covariance in the needlet coefficients domain;  $A_2$ ,  $A_3$  and  $A_4$  include the off-diagonal covariances, for which the localization of the needlets is crucial. In  $A_2$ , the “near-disjointness” of the supports of  $\psi_k$  and  $\psi_{k'}$  is only considered sufficiently far away from  $\xi_k$  and  $\xi_{k'}$ , while in  $A_3$  it is considered around  $\xi_k$  (or equivalently  $\xi_{k'}$ ). Finally,  $A_4$  depends on the regularity of  $p \mapsto \sigma_p^2$ .

We shall use the fact that :  $\forall x, C \in \mathbb{R}_+, \quad \frac{x^2}{(C+x^2)^2} \leq \min \left\{ \frac{x^2}{C^2}, \frac{1}{4C}, \frac{1}{x^2} \right\}$ .

With Cauchy-Schwartz inequality,  $W \leq 1$  and the properties of the cubature points (including Lemma 7.2) we have

$$\begin{aligned} A_1 &\leq \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) \leq \mathcal{B}^{-\kappa j}}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ \sum_{p=1}^N \lambda_p \sigma_p^2 \psi_k^2(\xi_p) \right\} \left\{ \sum_{p=1}^N \lambda_p \sigma_p^2 \psi_{k'}^2(\xi_p) \right\} \\ &= \sum_{k \in \mathcal{K}} \frac{n_k^2}{(C + n_k^2)^2} \sum_{\substack{k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) \leq \mathcal{B}^{-\kappa j}}} \frac{n_{k'}^2}{(C + n_{k'}^2)^2} \\ &\lesssim \frac{N^2}{C^2} \mathcal{B}^{-2\kappa j}. \end{aligned}$$

Thanks to Eq. (7.16) and Condition i)

$$\frac{A_1}{C^2 S^2} \lesssim N^2 \frac{(1+r)^4}{(\text{Card } V)^2} \mathcal{B}^{-2\kappa j} \asymp \mathcal{B}^{2j(2-\kappa_1-\kappa)} \xrightarrow{j \rightarrow \infty} 0$$

provided that

$$\kappa > 2 - \kappa_1 \quad . \quad (7.18a)$$

To give upper bounds on  $A_2$  and  $A_3$ , we use Proposition 7.3. Define

$$D(k, k') := \left\{ p : d(\xi_k, \xi_p) > \frac{1}{2} \mathcal{B}^{-\kappa j} \text{ and } d(\xi_{k'}, \xi_p) > \frac{1}{2} \mathcal{B}^{-\kappa j} \right\}.$$

Write

$$\begin{aligned} A_2 &\leq \sum_{k, k' \in \mathcal{K}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \\ &\quad \times \left\{ \sum_{p \in D(k, k')} \lambda_p W_p^2 \sigma_p^2 \psi_k(\xi_p) \psi_{k'}(\xi_p) \right\}^2 \\ &\lesssim \sum_{k, k' \in \mathcal{K}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \\ &\quad \times \left\{ \mathcal{B}^{2j} \sum_{p \in D(k, k')} \frac{\lambda_p W_p^2 \sigma_p^2}{(1 + \mathcal{B}^j d(\xi_p, \xi_k))^M (1 + \mathcal{B}^j d(\xi_p, \xi_{k'}))^M} \right\}^2 \\ &\leq \sum_{k, k' \in \mathcal{K}} (C + n_k^2)^{-2} (C + n_{k'}^2)^{-2} \left\{ \frac{\mathcal{B}^{2j}}{(1 + \frac{1}{2} \mathcal{B}^{(1-\kappa)j})^M} \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \right\}^2 \\ &\leq S^2 \left( \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \right)^2 \mathcal{B}^{2j(2-(1-\kappa)M)} \end{aligned}$$

so that, with Eq. (7.16) and Condition ii),

$$\frac{A_2}{C^2 S^2} \lesssim \left( \sum_{p=1}^N \lambda_p \sigma_p^2 \right)^2 \mathcal{B}^{2j(\alpha-(1-\kappa)M)} \asymp \mathcal{B}^{2j(\kappa_2+\alpha-(1-\kappa)M)} \xrightarrow{j \rightarrow \infty} 0$$

provided that

$$M > \frac{\kappa_2 + \alpha}{1 - \kappa} \quad . \quad (7.18b)$$

Remark now that if  $d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}$  and  $d(\xi, \xi_k) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}$  then  $d(\xi, \xi_{k'}) \geq \frac{1}{2} d(\xi_k, \xi_{k'})$  so that  $\frac{1}{(1 + \mathcal{B}^j d(\xi, \xi_k))(1 + \mathcal{B}^j d(\xi, \xi_{k'}))} \gtrsim \frac{1}{1 + \mathcal{B}^j d(\xi_k, \xi_{k'})}$ . Thus

$$\begin{aligned} A_3 &\lesssim \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}}} \frac{n_k^4}{(C + n_k^2)^2 (C + n_{k'}^2)^2} \sum_{p: d(\xi_k, \xi_p) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}} \lambda_p |\psi_k(\xi_p) \psi_{k'}(\xi_p)|^2 \\ &\lesssim \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \mathcal{B}^{-\kappa j}}} \frac{n_k^4}{(C + n_k^2)^2 (C + n_{k'}^2)^2} \mathcal{B}^{4j} \sum_{p: d(\xi_k, \xi_p) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}} \frac{\lambda_p}{(1 + \mathcal{B}^j d(\xi_k, \xi_{k'}))^2} \\ &\lesssim \sum_{k, k' \in \mathcal{K}} \frac{n_k^4}{(C + n_k^2)^2 (C + n_{k'}^2)^2} \mathcal{B}^{4j} \mathcal{B}^{-j(1-\kappa)2M} \mathcal{B}^{-2\kappa j} \\ &\leq \frac{N^2}{C^2} \mathcal{B}^{2j(2-(1-\kappa)M)} \end{aligned}$$

and with Condition i),

$$\frac{A_3}{C^2 S^2} \lesssim \frac{(1+r)^4}{(\text{Card } V)^2} N^2 \mathcal{B}^{2j(2-(1-\kappa)M)} \asymp \mathcal{B}^{2j(4-\kappa_1-(1-\kappa)M)} \xrightarrow{j \rightarrow \infty} 0$$

provided that

$$M > \frac{4 - \kappa_1}{1 - \kappa} . \quad (7.18c)$$

For the last term  $A_4$ , we use Condition ii), the same arguments as for  $A_3$ , and Lemma 7.3 which takes advantage of the localization of  $\psi_k$  around point  $\xi_k$  to compare the local noise level in direct and needlet domains. The estimate depends on the regularity  $\varrho$  of  $\sigma^2$ .

**Lemma 7.3.** Define  $\varrho_j := \max_{1 \leq k < k' \leq N_j} \frac{|(W_{jk}\sigma_{jk})^2 - (W_{jk'}\sigma_{jk'})^2|}{d(\xi_{jk}, \xi_{jk'})}$ . There exist a constant  $c > 0$  such that for all  $j \in \mathbb{N}$ ,  $\epsilon > 0$ , and  $k, k' \in \{1, \dots, N_j\}$  with  $d(\xi_k, \xi_{k'}) \leq \epsilon$ ,

$$|(W_{jk'}\sigma_{jk'})^2 - (n_{j,k})^2| \leq c \left\{ (n_{j,k})^2 \left(1 - \mathcal{B}^{\frac{1-M}{1+M}j}\right)^{-1} + \varrho_j \left(\epsilon + \mathcal{B}^{\frac{1-M}{1+M}j}\right) \right\}.$$

The proof of Lemma 7.3 is postponed to the end of the Section. For

$$M \geq \frac{1 + \kappa}{1 - \kappa} \quad (7.18d)$$

and using Lemma 7.3 with  $\epsilon = \frac{1}{2} \mathcal{B}^{-\kappa j}$  we get

$$\begin{aligned} A_4 &\lesssim \left(1 - \mathcal{B}^{\frac{1-M}{1+M}j}\right)^{-2} A_3 \\ &\quad + \varrho^2 \left(\frac{1}{2} \mathcal{B}^{-\kappa j} + \mathcal{B}^{\frac{1-M}{1+M}j}\right)^2 \sum_{\substack{k, k' \in \mathcal{K} \\ d(\xi_k, \xi_{k'}) > \frac{1}{2} \mathcal{B}^{-\kappa j}}} \frac{\sum_{p: d(\xi_k, \xi_p) \leq \frac{1}{2} \mathcal{B}^{-\kappa j}} \lambda_p |\psi_k(\xi_p) \psi_{k'}(\xi_p)|^2}{(C + n_k^2)^2 (C + n_{k'}^2)^2} \\ &\lesssim A_3 + \varrho^2 \left(\frac{1}{2} \mathcal{B}^{-\kappa j} + \mathcal{B}^{\frac{1-M}{1+M}j}\right)^2 \sum_{k, k' \in \mathcal{K}} \frac{1}{(C + n_k^2)^2 (C + n_{k'}^2)^2} \mathcal{B}^{4j} \mathcal{B}^{-j(1-\kappa)2M} \mathcal{B}^{-2\kappa j} \\ &\lesssim A_3 \left(1 + \varrho^2 \mathcal{B}^{-2\kappa j} \frac{1}{C^2}\right) \end{aligned}$$

as  $\frac{1-M}{1+M} < -\kappa$ . Note now that under ii) and using (7.1c) and (7.1e), we have

$$\varrho_j \underset{j \rightarrow \infty}{=} O(\mathcal{B}^{(\kappa_2+3)j});$$

and

$$\frac{A_4}{C^2 S^2} \lesssim \frac{A_3}{C^2 S^2} + \mathcal{B}^{2j(2-\kappa_1+\kappa_2+3+\alpha-(1-\kappa)M)} \xrightarrow{j \rightarrow \infty} 0$$

provided that

$$M > \frac{5 + \alpha - \kappa_1 + \kappa_2}{1 - \kappa} . \quad (7.18e)$$

It remains to see that there exists a  $\kappa \in (0, 1)$  satisfying simultaneously the conditions (7.18a) to (7.18e). This is the case if  $M > \max \left\{ \frac{\kappa_2 + \alpha}{\kappa_1 - 1} ; \frac{4 - \kappa_1}{\kappa_1 - 1} ; \frac{5 + \alpha - \kappa_1 + \kappa_2}{\kappa_1 - 1} \right\}$ . It suffices to take  $\kappa \in (2 - \kappa_1, 1)$  sufficiently near to  $2 - \kappa_1$  (recall that  $\kappa_1 > 1$ ). Finally, remark that any  $\kappa_1$  satisfying i) is bounded by 2, then the above condition on  $M$  reduces to the one mentioned in the statement of the Theorem. This concludes the proof of (7.14) and Theorem 7.2.

### Proof of Lemma 7.3

First, for any  $\delta > 0$ , using (7.1b), (7.1e), (7.8),  $0 \leq W \leq 1$  and the definition of  $\varrho$

$$\begin{aligned}
|W_k^2 \sigma_k^2 - n_k^2| &= \left| W_k^2 \sigma_k^2 - \frac{1}{\lambda_k} \sum_{p=1}^N \lambda_p^2 W_p^2 \sigma_p^2 \psi_k^2(\xi_p) \right| \\
&\leq \frac{1}{\lambda_k} \sum_{p=1}^N \lambda_p^2 |W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2| \psi_k^2(\xi_p) + W_k^2 \sigma_k^2 \left| 1 - \frac{1}{\lambda_k} \sum_{p=1}^N W_p^2 \lambda_p^2 \psi_k^2(\xi_p) \right| \\
&\lesssim \sum_{p: d(\xi_p, \xi_k) \leq \delta} \lambda_p |W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2| \psi_k^2(\xi_p) + \sum_{p: d(\xi_p, \xi_k) > \delta} \lambda_p |W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2| \psi_k^2(\xi_p) \\
&\quad + W_k^2 \sigma_k^2 \left( 1 + \sum_{p=1}^N \lambda_p \psi_k^2(\xi_p) \right) \\
&\lesssim \varrho \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \sum_{p: d(\xi_p, \xi_k) > \delta} \lambda_p |W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2| + W_k^2 \sigma_k^2 \\
&\lesssim \varrho \left( \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) + W_k^2 \sigma_k^2 \quad .
\end{aligned} \tag{7.19}$$

Second

$$\begin{aligned}
n_k^2 &\asymp \sum_{p=1}^N \lambda_p W_p^2 \sigma_p^2 \psi_k^2(\xi_p) = W_k^2 \sigma_k^2 \sum_{p=1}^N \lambda_p \psi_k^2(\xi_p) - \sum_{p=1}^N \lambda_p (W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2) \psi_k^2(\xi_p) \\
&= W_k^2 \sigma_k^2 \sum_{p=1}^N \lambda_p \psi_k^2(\xi_p) - \sum_{p: d(\xi_p, \xi_k) \leq \delta} (W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2) \psi_k^2(\xi_p) - \sum_{p: d(\xi_p, \xi_k) > \delta} (W_k^2 \sigma_k^2 - W_p^2 \sigma_p^2) \psi_k^2(\xi_p) \\
&\gtrsim W_k^2 \sigma_k^2 \left( 1 - \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) - \varrho \left( \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) \quad .
\end{aligned} \tag{7.20}$$

Combining Eqs. (7.19) and (7.20), one gets, for any  $k'$  with  $d(\xi_k, \xi_{k'}) \leq \epsilon$

$$\begin{aligned}
|W_{k'}^2 \sigma_{k'}^2 - n_k^2| &\leq |W_{k'}^2 \sigma_{k'}^2 - W_k^2 \sigma_k^2| + |W_k^2 \sigma_k^2 - n_k^2| \\
&\lesssim \varrho \epsilon + \varrho \left( \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) \\
&\quad + \left( n_k^2 + \varrho \left( \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) \right) \left( 1 - \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right)^{-1} \\
&= n_k^2 \left( 1 - \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right)^{-1} \\
&\quad + \varrho \left[ \epsilon + \left( \delta + \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right) \left( 1 + \left( 1 - \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} \right)^{-1} \right) \right] \quad .
\end{aligned}$$

Particularizing this last expression with  $\delta = \mathcal{B}^{\frac{1-M}{1+M}j}$  leads to Lemma 7.3, since we have, for  $M \geq 3$ ,  $j \geq 1$  and this  $\delta : \frac{\mathcal{B}^j}{(1 + \mathcal{B}^j \delta)^M} < \frac{1}{2}$ .

## Chapitre 8

# A full sky, low foreground, high resolution CMB map from WMAP

Ce chapitre reproduit l'article du même nom, J. Delabrouille, J.-F. Cardoso, M. Le Jeune, M. Betoule, G. Faÿ, F. Guillaoux, *Astronomy & Astrophysics* (à paraître). Seules certaines notations ont été modifiées pour la cohérence du manuscript.

**Abstract.** The WMAP satellite has made available high quality maps of the sky in five frequency bands ranging from 22 to 94 GHz, with the main scientific objective of studying the anisotropies of the Cosmic Microwave Background (CMB). These maps, however, contain a mixture of emissions from various astrophysical origins, superimposed on CMB emission. The objective of the present work is to make a high resolution CMB map in which contamination by such galactic and extra-galactic foregrounds, as well as by instrumental noise, is as low as possible. The method used is an implementation of a constrained linear combination of the channels with minimum error variance, and of Wiener filtering, on a frame of spherical wavelets called needlets, allowing localised filtering in both pixel space and harmonic space. We obtain a low contamination low noise CMB map at the resolution of the WMAP W channel, which can be used for a range of scientific studies. We obtain also a Wiener-filtered version with minimal integrated error. The resulting CMB maps offer significantly better rejection of galactic foregrounds than previous CMB maps from WMAP data. They can be considered as the most precise full-sky CMB temperature maps to-date.

### Sommaire

8.1	Introduction . . . . .	153
8.2	General considerations . . . . .	155
8.2.1	Requirements . . . . .	155
8.2.2	Evaluation and comparison of reconstructed CMB temperature . . . . .	157
8.2.3	Methods . . . . .	159
8.3	Evaluation and comparison of available maps . . . . .	160
8.3.1	Available maps . . . . .	162
8.3.2	Map comparison . . . . .	167
8.4	The ILC needlet method . . . . .	168
8.4.1	The choice of the ILC . . . . .	168
8.4.2	The ILC bias . . . . .	170

8.4.3	Needlets . . . . .	170
8.4.4	The method . . . . .	172
8.4.5	Point source subtraction . . . . .	172
8.4.6	Blanking of compact regions . . . . .	173
8.4.7	Needlet decomposition . . . . .	174
8.4.8	ILC implementation on needlet coefficients . . . . .	174
8.4.9	Full map reconstruction . . . . .	175
8.4.10	Final Wiener filtering . . . . .	175
8.4.11	Noise level estimate . . . . .	176
8.4.12	Bias estimates . . . . .	176
8.4.13	Practical implementation . . . . .	176
8.5	Simulations . . . . .	<b>177</b>
8.5.1	The simulated data . . . . .	177
8.5.2	Results . . . . .	177
8.5.3	Bias . . . . .	179
8.6	Application to WMAP data . . . . .	<b>179</b>
8.6.1	ILC Result . . . . .	179
8.7	Discussion . . . . .	<b>181</b>
8.7.1	Comparison with other maps . . . . .	181
8.7.2	Map characterisation . . . . .	186
8.7.3	Final comments . . . . .	190
8.8	Conclusions . . . . .	<b>191</b>
8.9	Derivation of the ILC bias . . . . .	<b>192</b>
8.9.1	First order expansion of the reconstruction error . . . . .	193
8.9.2	A geometric interpretation . . . . .	196
8.9.3	Comment on coefficient independence . . . . .	198

---

## 8.1 Introduction

The WMAP satellite is one of the most successful experiments dedicated to mapping the Cosmic Microwave Background (CMB). The all-sky maps obtained in the WMAP five frequency bands, in temperature and polarisation, offer the best data set to-date for making a sensitive all-sky map of the CMB anisotropies.

The CMB, however, is not the only source of emission at WMAP frequencies. Diffuse galactic emission from several processes contaminates the maps with an amplitude roughly proportional to the cosecant of the galactic latitude, compromising the observation of the CMB close to the galactic plane. In addition, a background of radio and infrared compact sources, galactic or extra-galactic, contributes to the total emission even at high galactic latitudes. Component separation consists in separating one or more of these sources of emissions from the others in the data.

One of the main objectives of CMB experiments is the measurement of the CMB angular power spectrum  $C_\ell$  which, with the assumption of statistical isotropy, describes the second order distribution of the CMB and can be estimated on a fraction of the sky. For this reason, many ground-based and balloon-borne experiments have concentrated their observations in “clean” regions of the sky, where galactic emission is low enough to impact negligibly the observations. For power spectrum estimation

from full-sky observations, a safe approach consists in masking regions at low galactic latitude, and estimating power spectra on the cleanest regions of the sky. The impact of extragalactic point sources (evenly spread on the sky) on power spectrum estimates can be evaluated and corrected for using ancillary data (catalogues of known point sources and priors on the statistical distribution of sources).

Besides the power spectrum, the CMB map itself is interesting for several additional purposes :

- As a CMB template, to be subtracted from millimetre–wave observations when the scientific focus is on other emissions, or to be used for the calibration of other instruments ;
- To assess the statistical isotropy of the CMB and check the homogeneity and isotropy of the Universe on the largest scales ;
- To search for signatures of non-trivial topology, as that of a multi-connected universe [52, 61, 113] ;
- To search for correlations of the CMB map with other emissions [115, 81, 60, 62, 121, 109, 123] ;
- To search for signatures of non Gaussianity in the CMB [112, 142, 64, 59, 108, 140, 107, 109, 147].

Several CMB maps obtained with the WMAP data are available for such research projects. The WMAP team has released part–sky foreground–reduced maps in the Q, V and W bands, and maps obtained by an Internal Linear Combination (ILC) of all WMAP channels [86, 82]. Tegmark et al. [139] have produced CMB maps with WMAP one year data, and subsequently with three year data, based on an harmonic space ILC method. Eriksen et al. [75] have an alternate version of the ILC CMB map at 1 degree resolution on one year data. Eriksen et al. [76] use a Gibbs sampler to draw 100 realisations of the CMB under the posterior likelihood of a model of CMB and foregrounds ; their estimated CMB map is the average of these realisations for three year data. On three year data again, Park et al. [118] use an ILC technique on 400 different pixel ensembles, selected by the spectral index of the foreground emission as estimated by the WMAP team using a Maximum Entropy Method (MEM). More authors have addressed component separation on WMAP three year data and produced versions of a “clean” CMB map [103, 128, 57]. More recently, Kim et al. [94] have obtained a CMB map from WMAP five year data.

All available maps suffer from limitations, some of which result from specific choices in the way the CMB map is produced. Several of these maps, for instance, do not fully exploit the resolution of the original observations. Some focus on cleaning the CMB from foregrounds at high galactic latitude, and are significantly contaminated by foregrounds in the galactic plane. Some are not full sky CMB maps. Finally, not all of the available maps have well characterised noise and effective beam. All these limitations impact their usefulness for accurate CMB science.

In this paper, we address the problem of making a CMB temperature map which has the following properties :

- being full sky ;
- being as close as possible to the true CMB (minimum variance of the error) everywhere on the sky, and on all scales ;
- having the best resolution possible ;
- having well-characterised beam and noise.

In the following, we first review these requirements and their impact on a CMB cleaning strategy (Section 8.2). We then review and compare available maps in Section 8.3. In Section 8.4, we describe and explain our ILC needlet method. The approach is tested and validated on realistic simulated data sets (Section 8.5) before applying it to WMAP data (Section 8.6). We then compare our CMB maps to the other existing maps, discuss the results, and conclude.

This paper considers only *temperature* maps.

## 8.2 General considerations

### 8.2.1 Requirements

We start with a review of the requirements above, and on the implication on the method to be used.

#### Full sky ?

WMAP data are full sky. We wish to devise a method which allows recovering an estimate of CMB emission everywhere, including in the galactic plane, and even (as much as possible) in the galactic centre as well as in pixels strongly contaminated by compact sources.

The large scale correlation properties of the CMB make it possible to estimate the CMB emission even in unobserved regions, by some kind of interpolation. This is incidentally what is obtained with the Gibbs sampling technique of Eriksen et al. [76]. Equivalent in spirit although quite different in implementation is the use of an “inpainting” method as that of Elad et al. [12]. Such interpolation methods alone are not fully satisfactory, as they discard information. In particular they do not allow recovering small scale CMB features in the mask. This is obvious, for instance, in the Gibbs-sampling average map of Eriksen et al. [76].

At the opposite, one may try to separate components in the galactic plane independently of what is done at higher galactic latitudes, since levels and properties of foreground emission depend strongly on sky direction. Hinshaw et al. [86] and Tegmark et al. [139] divide the sky into several independent regions, perform component separation independently in these regions, and then make a composite map by stitching together these independent solutions. Such approaches discard information (zone-to-zone correlations) and require careful treatment at the zone borders to avoid discontinuities and ringing.

A good method should perform well on both counts : localised processing and full exploitation of large scale correlations of the CMB and of galactic foregrounds. This can be achieved with a spherical wavelet or *needlet* analysis of the maps (using the tools developed in Marinucci et al. [105] and Guilloux et al. [25]), which is our approach in the present work.

#### Minimum variance ?

Recovery of a CMB map can be conducted following various objectives, quantified by different “figures of merit”. In this work, we choose, as most authors, to minimise the variance of the difference between true and recovered CMB (this is the *overall* error ; it includes additive noise, foregrounds, and even multiplicative errors affecting the CMB itself).

This choice alone does not fully characterise the method to be used. The best theoretical solution also depends on the model of the data. An overview of existing methods can be found in the review by Delabrouille and Cardoso [67].

In this paper, contrarily to other approaches which rely heavily on the structure of the data as described by a model, for instance a noisy linear mixture of independent components [69], we assume as little as possible about the foregrounds and the noise. In fact, we assume nothing except the following :



- The WMAP data are well calibrated with known beam in each channel ;
  - The instrumental noise in all WMAP maps is close to being Gaussian and uncorrelated ; its pixel-dependent level is approximately known ;
  - The CMB anisotropies emission law is known to be the derivative with respect to temperature of a  $T=2.725\text{K}$  black-body ;
  - To first order, the template of CMB anisotropies is well represented by a Gaussian stationary random field, the spectrum of which is given by the WMAP best fit (as will be seen later-on, this last assumption is needed only to derive the Wiener filter ; it is not necessary for our needlet ILC map).
- These assumptions lead us to consider an “Internal Linear Combination” (ILC) method, followed by a Wiener filter to minimise the error integrated over all scales.

### **Best resolution ?**

The WMAP data comes in five frequency channels with varying resolution. To make the best of the data, we need a method which uses the smallest scale information from the W band, and information from additional bands (V, then Q, then Ka, and finally K) at increasingly larger scales. Multi-scale tools are well suited for this purpose.

The “best possible resolution” is not a well defined concept (and not necessarily the resolution of the W band). Indeed, there is a conflict between best possible resolution and minimum variance, as one can smooth or deconvolve a map arbitrarily in harmonic space, reducing or increasing the total noise variance in the process. Here, we make a map at the resolution of the W channel over the full sky, leaving open the option to filter this map if needed to reduce the noise – or deconvolve it for better angular resolution. Note that additional global filtering or deconvolution does not change the signal to noise per mode (only the integrated S/N).

The minimum variance map is obtained with a Wiener filter, which smooths the map depending on the signal to noise ratio. As the noise and the contaminants are inhomogeneous, there is a strong motivation for the smoothing to depend on the location on the sky (the optimal solution to the resolution–variance trade–off depending on the contamination level, which is local). If we relax the constraint about beam homogeneity, again, spherical needlets offer a natural way to obtain such location–dependent smoothing. In the present, however, in order to preserve the constancy of the effective beam over the sky, we implement the Wiener filter in harmonic space.

### **Accurate characterisation ?**

A fully accurate characterisation of the beam and noise is not straightforward, in particular because of the limited knowledge about the original frequency maps, which automatically propagates into the final CMB map. This work makes several approximations about beams and noise. Beams are assumed symmetric and therefore described by the  $B_\ell$  transfer functions provided by the WMAP team. The instrumental noise is assumed uncorrelated, although non stationary. Analytical analyses and Monte-Carlo simulations are used to characterise the residual noise of the final map, as well as to estimate the contribution of residual foregrounds, and biases if any. This will be detailed further later-on.

## Noise

Throughout this paper, the term “noise” typically stands for all sources of additive error, i.e. instrumental noise *and* foregrounds.

### 8.2.2 Evaluation and comparison of reconstructed CMB temperature

We briefly discuss here the tools used for characterising and comparing CMB temperature maps.

#### Map description

A pixelised map is fully characterised by the specification of :

- A set of temperature values  $y_p$  in a number of pixels (here indexed by  $p$ ) ;
- The effective beam at each pixel  $p$ , which in the most general case is a function  $B_{p,p'}$  ;
- The noise  $n_p$ , the statistical properties of which, in the Gaussian case, are fully described by a noise covariance matrix  $N_{p,p'}$ .

The map value  $y_p$  is then linked to the true signal value  $s_p$  by :

$$y_p = \sum_{p'} B_{p,p'} s_{p'} + n_p \quad (8.1)$$

The full characterisation of a given CMB map requires the specification of the additive noise  $n_p$  and of the response  $B_{p,p'}$ . When the beam is stationary over the sky and symmetric, which we assume in this work, it is fully specified by the coefficients  $B_\ell$  of the expansion of the beam in Legendre polynomials.

#### Assumptions

Throughout this paper, the beam is assumed symmetric. Although this is an approximation, most pixels of the WMAP map are “visited” by any particular detector through a wide range of intersecting scans. The average beam in that pixel then is an average of the physical beam over many orientations, which makes the symmetry assumption reasonable.

In addition, in absence of any specific localised processing, the beam is assumed to be invariant over the sky.

With the above two assumptions, the effect of beam convolution is best represented in harmonic space, with a multiplicative coefficient  $B_\ell$ , independent of  $m$ , applied to the harmonic coefficients  $a_{\ell m}$  of the map. We assume perfect beam knowledge as well as perfect calibration, so that no multiplicative uncertainty is attached to the map description (the beam integral, approximated as  $\sum_{p'} B_{p,p'}$ , is equal to unity independently of  $p$ , or, equivalently, the value of  $B_\ell$  for  $\ell = 0$  is assumed to be exactly unity).

The noise  $n_p$  of the original WMAP maps, for each frequency channel and each differencing assembly, is assumed uncorrelated from pixel to pixel, i.e.  $N_{p,p'} = \langle n_p n_{p'} \rangle = \delta_{pp'} \sigma_p^2$ . The variance  $\sigma_p^2$  is pixel dependent because of uneven sky coverage. Noise is non-stationary, but assumed to be Gaussian distributed.

## Comparison of maps at different resolution

The comparison of CMB maps is meaningful only if the maps are at the same resolution. As long as the beam transfer function does not vanish at any useful  $\ell$  (which is always the case for Gaussian beams), the resolution of any map can be changed to anything else by harmonic transform and multiplication by the ratio of the beam transfer functions. In the present work, we approximate WMAP beams by their symmetric fit (i.e. azimuthally averaged beams), which is justified by the large range of scanning directions at any point in the sky.

This convolution – deconvolution property is widely used throughout this paper.

## Masking

We define “tapered” regions of the sky for map comparison at varying galactic latitude. In particular, we define a Low Galactic Latitude (LGL) region and a complementary High Galactic Latitude (HGL) region. The LGL region, used to evaluate results in the galactic plane, cuts completely all data above 30 degrees galactic latitude (and below  $-30$  degrees), and has a 15 degree transition zone with a cosine square shape. All pixels at absolute galactic latitudes below 15 degrees are kept with a coefficient of 1. The HGL region is the complementary, i.e.  $\text{HGL} = 1 - \text{LGL}$ . These “tapered” regions allow the computation of local power spectra with negligible spectral leakage of large scale power into small scales.

## Power spectra comparison

For a given beam (i.e. multiplicative response as a function of  $\ell$ ), the comparison of the total map power as a function of  $\ell$  (i.e. of the power spectra of the maps), is a direct figure of merit. The lower the power spectrum, the better the map.

Indeed, for fixed  $b_\ell$ , two distinct maps contain the exact same contribution from CMB, but different noise levels. CMB and noise being assumed to be uncorrelated, power spectra are the sum of a CMB term ( $b_\ell^2 C_\ell$ ) and a noise term ( $N_\ell$ ). The CMB term being the same for both spectra, any difference in total power spectrum can be interpreted as a difference in noise level.

Power spectra are computed independently for different regions of the sky (e.g. inside or outside the galactic plane). To minimise aliasing due to sharp cuts, we use masks with smooth transitions, defining the LGL and HGL regions described above.

The power spectrum of a map is evaluated as :

$$\widehat{C}_\ell = \frac{1}{(2\ell + 1)\alpha} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \quad (8.2)$$

$\alpha$  is a normalising factor computed as the average value of the squared masking coefficients.<sup>1</sup>

Power spectra estimated directly in this way from a masked sky are unreliable for modes corresponding to angular sizes larger than the typical size of the zone of the sky retained for computation.

---

<sup>1</sup>The masking coefficient is simply 1 in regions kept for power spectrum computations, 0 in regions masked, and between 0 and 1 in the transitions.

### 8.2.3 Methods

We now give a brief introduction to the two main methods used in this paper (ILC and Wiener filtering). Many other methods exist for CMB cleaning (or component separation in general), which assume varying degrees of prior knowledge about sky emission, and model the data in different ways. These methods are not discussed nor used in this paper. For a review, see Delabrouille and Cardoso [67].

#### The ILC

The data are modelled as

$$\mathbf{x} = \mathbf{a}s + \mathbf{n} \quad (8.3)$$

where  $\mathbf{x}$  is the vector of observations (e.g. five maps),  $\mathbf{a}$  is the response to the CMB for all observations (e.g. a vector with 5 entries equal to 1 if WMAP data only are considered) and  $\mathbf{n}$  is the noise. Here it is assumed that all observations are at the same resolution.

The ILC provides an estimator  $\hat{s}_{\text{ILC}}$  of  $s$  as follows :

$$\hat{s}_{\text{ILC}} = \frac{\mathbf{a}^t \widehat{\mathbf{R}}^{-1}}{\mathbf{a}^t \widehat{\mathbf{R}}^{-1} \mathbf{a}} \mathbf{x} \quad (8.4)$$

where  $\widehat{\mathbf{R}}$  is the empirical covariance matrix of the observations (e.g., a  $5 \times 5$  matrix when 5 channels are considered). The biggest concern with the ILC (and the source of all problems of bias) is that this empirical covariance matrix has to be estimated on the data themselves. This will be discussed further in section 8.4

Note that the ILC solution of Equation 8.4 is the linear filter which minimises the total variance of the output map, under the condition that the filter has unit response to the signal of interest (the signal with emission law given by vector  $\mathbf{a}$ ).

The details of the method for its implementation in the context of this work are further discussed in Section 8.4.

#### Wiener filtering

Given a single CMB map of known beam (assumed to be constant over the sky), it is possible to minimise the contamination by noise and foregrounds by (one-dimensional) Wiener filtering. The data is modelled as  $x = s + n$ , where now  $x$  is a single map,  $s$  the true CMB and  $n$  the noise. The Wiener filter gives to individual “modes” a weight proportional to the fraction of signal power in that mode, i.e.

$$\hat{s}_{\ell m} = \frac{B_\ell^2 C_\ell}{B_\ell^2 C_\ell + N_\ell} x_{\ell m} \quad (8.5)$$

where  $B_\ell^2 C_\ell$  and  $N_\ell$  are the power spectra of the (smoothed) CMB and of the noise (including smoothed foregrounds or foreground residuals) respectively, and  $x_{\ell m}$  is the original noisy CMB map. It

should be noted that if the CMB and the noise are uncorrelated, then  $B_\ell^2 C_\ell + N_\ell = X_\ell$  is the power spectrum of the map  $x_{\ell m}$ , and the Wiener filter can be estimated directly using only a prior on the CMB power spectrum (assuming  $C_\ell$  is known), and estimating  $X_\ell$  on the map itself.

Wiener filtering in harmonic space minimises the variance of the error in the map if signal and noise are Gaussian and stationary. For non-stationary contaminants, the Wiener filter (8.5) is still meaningful, but is no longer optimal. Here, the Wiener filtering is applied after the ILC. Foreground contamination, strongly reduced by the ILC, is no longer dominant. Instead, instrumental noise is the major source of error on all scales where the Wiener filter is useful (departs from 1). An harmonic space implementation of the Wiener filter disregards this non-stationarity, applying the same coefficient to a given scale, whether it is in a region of deeper integration or not.

Better efficiency may be obtained by an implementation in another domain than the harmonic space (e.g. needlets), but this is not investigated further in this paper, as the price of pay is non stationary equivalent beam.

### 8.3 Evaluation and comparison of available maps

Before describing how to make yet another CMB map from WMAP data, we review the existing maps and evaluate in which respect they can be improved.

F . 8.1 – Available CMB maps.

NAME	resolution	data used	Ref.	URL
WILC1	1°	1-yr	[54]	<a href="http://lambda.gsfc.nasa.gov/product/map/dr1/imaps_ILC.cfm">http://lambda.gsfc.nasa.gov/product/map/dr1/imaps_ILC.cfm</a>
TILC1	W channel	1-yr	[139]	<a href="http://space.mit.edu/home/tegmark/wmap/cleaned_map.fits">http://space.mit.edu/home/tegmark/wmap/cleaned_map.fits</a>
EILC1	1°	1-yr	[75]	<a href="http://www.astro.uio.no/~hke/cmbdata/WMAP_ILC_lagrange.fits">http://www.astro.uio.no/~hke/cmbdata/WMAP_ILC_lagrange.fits</a>
WILC3	1°	3-yr	[86]	<a href="http://lambda.gsfc.nasa.gov/product/map/dr2/ilc_map_get.cfm">http://lambda.gsfc.nasa.gov/product/map/dr2/ilc_map_get.cfm</a>
EGS3	3°	3-yr	[76]	<a href="http://www.astro.uio.no/~hke/gibbs_data/cmb_mean_stddev_WMAP3_n64_3deg.fits">http://www.astro.uio.no/~hke/gibbs_data/cmb_mean_stddev_WMAP3_n64_3deg.fits</a>
PILC3	1°	3-yr	[118]	<a href="http://newton.kias.re.kr/~parkc/CMB/SILC400/SILC400_bc.fits">http://newton.kias.re.kr/~parkc/CMB/SILC400/SILC400_bc.fits</a>
TILC3	W channel	3-yr		<a href="http://space.mit.edu/home/tegmark/wmap/cleaned3yr_map.fits">http://space.mit.edu/home/tegmark/wmap/cleaned3yr_map.fits</a>
WILC5	1°	5-yr	[82]	<a href="http://lambda.gsfc.nasa.gov/product/map/dr3/ilc_map_get.cfm">http://lambda.gsfc.nasa.gov/product/map/dr3/ilc_map_get.cfm</a>
KILC	1°	5-yr	[94]	<a href="http://www.nbi.dk/~jkim/hilc/">http://www.nbi.dk/~jkim/hilc/</a>

We start with a discussion of the existing methods, identifying for each of them strengths and weaknesses of the approach, and their foreseeable consequences. Available CMB maps obtained from WMAP data are compared in terms of resolution, of the estimated contamination by foregrounds and of noise level. In absence of an absolute reference, discrepancies between available maps are also evaluated. This comparison permits to estimate typical uncertainties, and to outline the “difficult regions” for CMB cleaning (which, unsurprisingly, are mostly located close to the galactic plane). We also look specifically for residuals of galactic contamination by comparing the power spectrum of the reconstructed CMB map at high and at low galactic latitudes. Significant discrepancies between the two are interpreted as indicative of a residue of foreground emission.

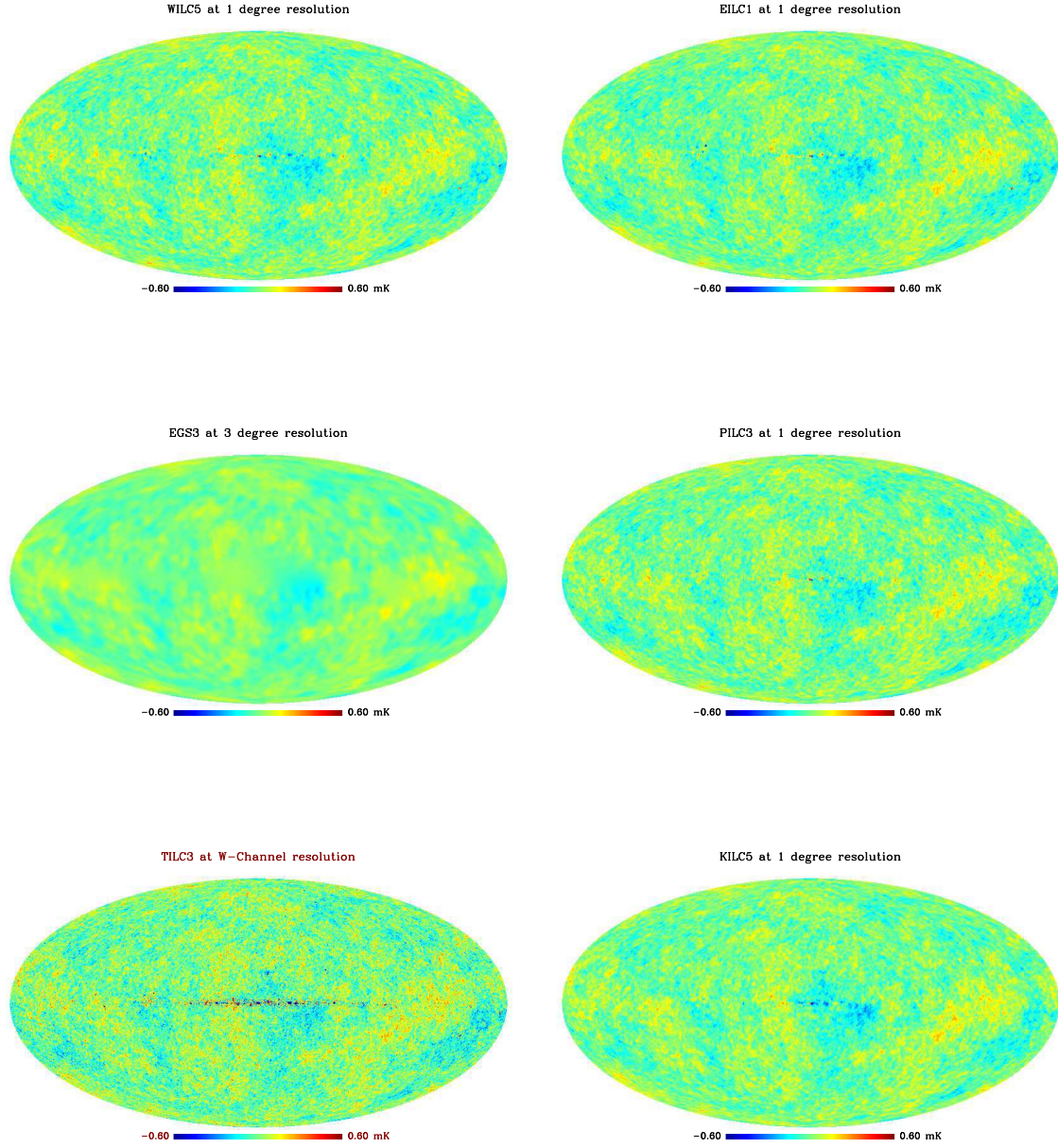


FIG. 8.2 – This figure shows a selection of presently available CMB maps from WMAP data. Contamination by galactic emission is visible in all of them at various levels, except for the 3-degree resolution EGS3 map (for which a galactic cut was applied and then filled by a plausible CMB extrapolated from higher galactic latitude data).

### 8.3.1 Available maps

#### The WMAP ILC

The ILC maps obtained by the WMAP team (denoted as WILC1, WILC3 and WILC5 hereafter, depending on whether they are obtained using one year, three year or five year data) are described in Bennett et al. [54], Hinshaw et al. [86], and Gold et al. [82] respectively.

For the three year and five year maps, the original frequency maps are smoothed to a common resolution of one degree. The sky is subdivided into 12 regions. One big region covers most of the sky at moderate to high galactic latitudes. The rest of the sky, concentrated around the galactic plane, is divided into 11 regions of varying galactic emission properties (amplitude and colour). An Internal Linear Combination is performed independently in each of these regions. A full sky composite map is obtained by co-adding the maps of the individual regions (with a  $\simeq 1$  degree transition between the zones to avoid sharp edge effects). Finally, a bias (unavoidable consequence of empirical CMB-foreground correlation) is estimated by Monte-Carlo simulations, and subtracted from the composite map, to yield the final CMB map.

The three year and five year ILC maps differ from the one year map in several respects. The most significant is the recognition of the existence of a bias, and the attempt at correcting it using simulations. The limitations of the maps include their resolution (one degree), and the use of small regions in the ILC, which is bound to cause more bias than necessary on large scales (comparable to patch sizes),<sup>2</sup> as well as edge effects. This results in discontinuities between regions, obvious for instance in the estimated bias map shown in Hinshaw et al. [86].

It is worth mentioning that in the method used by the WMAP team, the coefficients of the linear combination used over most of the sky (region 0, which corresponds to the largest part of the sky at high galactic latitudes and a few low galactic latitude patches, and region 1, in the galactic plane but away from the galactic center) are set using only a small subset of the data inside the Kp2 cut (where the galactic emission is the strongest). This choice favours the rejection of galactic contamination, at the price of sub-optimal weighting of the observations in regions where the error is dominated by noise. It also assumes that the emission laws and relative power of the different foregrounds are the same in these regions, which is a strong (and probably wrong) assumption.

Furthermore, the ILC weights are set by minimising the variance of the map at one degree resolution. Modes at higher  $\ell$  get very sub-optimal weighting, as they do not contribute significantly to the total variance of the one degree map. The K and Ka bands, in particular, contribute respectively to about 0.156 and  $-0.086$  for region 0 (the largest one) for WILC3. As a consequence, the final ILC map can not be meaningfully deconvolved to better resolution than about 1 degree (because this would blow-up small scale noise coming from the lowest frequency channels).

Finally, there is also an unsatisfactory degree of arbitrariness in the choice of the regions, which depend on priors about foreground emission, and are somewhat elongated across the galactic plane for no particular reason. Although none of these choices is unreasonable, the impact of this arbitrariness on the final map is difficult to evaluate.

For all these reasons, the WILC maps leaves considerable margin for improvement. We aim, in particular, at obtaining a CMB map with better angular resolution, and with a better handling of non-

---

<sup>2</sup>Small patches contain few modes, hence empirical correlations are stronger, as discussed in appendix A

stationarity and scale dependence of the contamination (foregrounds and noise).

### **The WMAP foreground-reduced maps**

For temperature power spectrum analysis, the WMAP team has used part-sky foreground-reduced maps. The processing for foreground removal for the three year and five year releases is described in Hinshaw et al. [86]. Model templates for galactic emission are fitted to the Q, V and W WMAP channels outside of the Kp2 mask. A linear combination of synchrotron, free-free and dust, based on this fit, is then subtracted from the full sky Q, V, W maps.

In this procedure, a first galactic template, supposed to correspond to a linear combination of synchrotron and free-free emission, is obtained from the difference between the K and Ka bands. This template is produced at one degree resolution. An additional free-free template is obtained from H $\alpha$  emission [79] corrected for dust extinction [54]. A dust template is obtained from model 8 of Finkbeiner et al. [80]. “Clean” Q, V and W maps are obtained by decorrelation of these templates from the original Q, V and W observations.

The main limitation of this approach is that the model used is insufficient to guarantee a good fit of the total foreground emission simultaneously inside and outside of the Kp2 mask. As a consequence, the maps produced are heavily contaminated by foregrounds in the galactic plane, the priority being given to higher galactic latitudes, with the objective of obtaining a part-sky high quality map on which high multipole CMB power spectra could be estimated reliably.

In addition, the maps are likely to depend significantly on the prior model assumed. Here, the WMAP team chooses dust model number 8 of Finkbeiner et al. [80], and also ignores the plausible existence of anomalous dust emission. The exact impact of these *a priori* decisions is difficult to evaluate. Anomalous dust emission will come into the foreground reduced maps chiefly as a contaminant in the K - Ka synchrotron template, and hence erroneous extrapolation of this template to higher frequencies. Assumptions about the emission laws of the template correlated to H-alpha, and of the dust template, if inaccurate, will also result in inaccurate subtraction of free-free and dust. These errors in the estimate of foregrounds are bound to leave, after template subtraction, foreground residuals in the foreground reduced maps.

As an additional drawback, we note that the method generates correlated noise in the foreground-reduced maps, originating either from K and Ka channel noise or from a background of radio sources (see, e.g., [91] and [146] for discussions on radio sources in WMAP data and their impact on the analysis of CMB observations with WMAP). Finally, on supra degree scales, the K and Ka bands, which are the most sensitive ones, are used only to subtract foregrounds, whereas in the cleanest regions of the sky they would be more usefully used to estimate the CMB emission.

For all these reasons, WMAP foreground-reduced maps are not good CMB maps according to the criteria listed in the introduction.

### **The ILC by Eriksen et al.**

Eriksen et al. [75] have obtained a CMB map at 1 degree resolution with another implementation of the ILC. The map, denoted here EILC1, uses only one year data.



An interesting remark from Eriksen et al. is that the amount of residual dust is high in the ILC maps – the method being able to subtract only about half of the dust present in the W band. At scales larger than 1 degree, this lack of performance is likely to be due to the part of dust emission uncorrelated to low frequency galactic foregrounds. On the smallest scales, the situation is worse, as the low frequency WMAP channels do not have the resolution to help remove small scale dust emission from W band observations.

For this reason, in the present work, we improve on dust removal by using, as an additional measurement, the IRIS 100 micron dust template obtained from a combination of DIRBE and IRAS maps [111]. As compared to the map of Eriksen et al., we also aim at better angular resolution – and, obviously, better sensitivity, achieved by using five year data sets.

### **The Gibbs-sampling map by Eriksen et al.**

Recently, Eriksen et al. [76] have produced a low resolution (3 degree) CMB map from WMAP three year data, using a Gibbs sampling technique to explore the likelihood of a parametric model of CMB and foreground emissions. A CMB map is obtained as the average of 100 sample CMB maps drawn each at random according to the posterior distribution of the model parameters given the observations.

The free parameters in the model are spherical harmonic coefficients  $a_{\ell m}$  of the CMB map, the CMB harmonic power spectrum  $C_\ell$ , monopole and dipole amplitudes in each WMAP band, the amplitude  $a(\nu)$  of a dust template in each band, and amplitudes  $f(p)$  and spectral indices  $\beta(p)$  of a low-frequency foreground component, for each map pixel  $p$ .

The model is constrained by fixing the dust template at 94 GHz according to Finkbeiner et al. [80], by a prior on the low-frequency foreground spectral index, assumed to be close to that of synchrotron ( $-3 \pm 0.3$ ), and by the constraint that the monopole and dipole coefficients are orthogonal to the (noise-weighted) pixel-averaged foreground spectrum.

In spite of a good fit of the assumed model to the data at high galactic latitudes, there are some strong limitations to the resulting CMB map, and hence to its usage :

- The result of the sampling is obtained assuming a parametric model of foreground emission. There is no possible way of validating the systematic errors due to mismodelling, except marginalising over all possible model skies. This would require a Monte-Carlo simulation which takes into account all uncertainties in the modelling, not only values of the parameters for a given parametric model, but also the choice of the parameter set to be used to model the foregrounds (varying the dust template according to uncertainties, assuming a different foreground model, etc...). This is not presently available ;
- The resulting map is at 3 degree resolution and HEALPix  $n_{\text{side}} = 64$ , considerably worse than WMAP can do ;
- The data sets are cut with the Kp2 mask. Although a CMB is recovered in the mask by the average of the sample maps, the effective resolution inside the mask is lower than in the rest of the sky. In some sense, the Gibbs sampling technique (as implemented here) allows to recover in the mask what is predictable from the outside map (assuming stationarity of the CMB anisotropy field). It allows only for a limited prediction of the CMB signal in the masked zone.

For all of these reasons, the Gibbs-sampling map of Eriksen et al. (hereafter EGS3) is not a good “best CMB” map according to our criteria.

### **The ILC by Park et al.**

Park et al. [118] provide their own version (hereafter the PILC3 map) of a one degree resolution CMB map obtained by an ILC on WMAP three year data. The originality of their approach lies in the fact that they cut the sky into 400 pixel ensembles, selected from a prior on their spectral properties. The 400 ensembles are defined from  $20 \times 20$  spectral index bins (20 for K-V spectral index, and 20 for V-W). This approach is motivated by the fact that ILC weights are expected to vary with varying foreground properties.

There is a weak point to this approach. The authors use, to define their pixel ‘bins’, the MEM solution derived by the WMAP team. If the MEM solution is wrong for a given pixel, that pixel will automatically be classified in the wrong pixel ensemble, and be weighted using the weights of the wrong population of pixels. To some extent then, this binning forces the result of the ILC to match the prior assumptions given by the MEM results. In turn, the MEM solution uses as a prior the result of the WMAP ILC, which is subtracted from the WMAP observations prior to using the MEM method to separate galactic foregrounds.

As a consequence, the connection of the CMB map of Park et al. to the original WMAP data is far from direct. The map is bound to bear the signature of any arbitrary choice made before, in particular the choice of WMAP ILC regions, the 3-component model for galactic emission, and the MEM priors. For instance, discontinuities at the boundaries between the 12 regions of the WMAP ILC are clearly visible in the map of K-V spectral index used by the authors, as well as their group index (see Figures 3a and 4a of their paper).

Park et al. then investigate the error in their reconstructed map by Monte-Carlo simulations. However, they use as an input galactic emission template the very model obtained by the MEM. This means that in the simulations, the spectral index maps are “exact”. Therefore, the simulations investigate accurately the errors only if the MEM solution is correct, which is not likely to be the case –at least not to the level of precision required for producing a CMB map useful for precision cosmology.

Our method described in Section 8.4 uses as little prior information as possible, and aims at better angular resolution than 1 degree.

### **The “clean” map of Tegmark et al.**

The approach of Tegmark et al. [139], on both one year and three year data, is the only work to date which aims at producing a CMB map with both full sky coverage and best possible resolution.

Tegmark et al. [139] have performed a foreground analysis of the WMAP one year maps, producing two high resolution maps of the CMB. The first one, the “clean” map (hereafter TILC1), is obtained by a variant of the ILC in which weights are allowed to vary as a function of  $\ell$ , and are computed independently in nine independent regions. The second, the “Wiener” map, is a Wiener-filtered version of the same map, in which the Wiener filter is applied in harmonic space, but independently in each zone.

This work by Tegmark et al. is an early attempt at finding linear combinations of the WMAP data which vary both in harmonic space and in pixel space. The pixel variation of the weights is made using zones which are defined according to the level of contamination by foregrounds, as computed from WMAP map differences W-V, V-Q, Q-K and K-Ka. The authors do not specify exactly how the

frequency range is divided into  $\ell$ -bands. Whereas the text seems to indicate that weights are computed independently for each  $\ell$ , figures hint that the weights are actually band-averaged, in 50 logarithmic bands subdividing the multipole range. It makes a difference for the ILC bias. For an ILC implemented  $\ell$  by  $\ell$ , the number of modes at a given  $\ell$  is  $2\ell + 1$ , and thus remains small for most of the useful range of scales, yielding significant bias (of order 10% at  $\ell = 25$ , and 1% at  $\ell = 250$ , for a full sky ILC; for part sky ILC, the bias is multiplied by something of order the inverse of the sky fraction). For band-averages, the effect would be less dramatic, because of the higher number of modes per individual ILC.

In the end, the authors obtain a CMB “clean” map with a “beam corresponding to the highest-resolution map band”, i.e. the beam of the W band.

The original paper describes the work done on the one year WMAP data. However maps for three year data are available on Max Tegmark’s web site (see Table 8.1). We use the three year map (TILC3) for comparison with our own solution.

Although the approach of Tegmark and collaborators is quite good at high galactic latitudes, we can see on Figure 8.2 that it performs poorly in the galactic plane. Also, the authors have not removed detected point sources from the WMAP data before making the ILC. As a result, their CMB map contains obvious point source residuals, for instance around galactic longitude  $305^\circ$  and latitude  $57^\circ$ , where a 5 mK peak can be seen.

Our method, although bearing some similarity with that of Tegmark and collaborators, aims at improving significantly the error characterisation, as well as the quality of foreground cleaning in the galactic plane.

### **The “Wiener” map of Tegmark et al.**

In addition to their TILC map, Tegmark et al. [139] publish a Wiener map (hereafter TW map), obtained from the TILC map by independent Wiener filtering in the 9 regions. This results in reduced integrated error in all regions, at the price of pixel-dependent extra smoothing. The consequence of this filtering is an effective zone-dependent beam.

Because of this extra smoothing, it is difficult to compare the TW map with other maps. The most meaningful figure of merit for the Wiener map, anyway, is the actual power of the error (output map minus true CMB). This is unavailable for any useful up-to-date real data set. Additional discussion about Wiener-filtered maps is deferred to Section 8.6.

### **The ILC map of Kim et al.**

More recently, Kim et al. [94] have made a CMB map from WMAP five year data, using an “harmonic” ILC method (KILC5 hereafter). Their method perform an ILC in the pixel domain but with pixel-dependent weights. The ILC weights are not constant over predefined zones on the sky but are computed as smooth weight maps defined in terms of an harmonic decomposition (hence the qualification of the method). More specifically, the weight maps are determined by minimising the total output CMB map variance with the constraint that these maps have no multipoles higher than  $\ell_{\text{cutoff}}$ . For stability reasons, the KILC5 map is obtained with  $\ell_{\text{cutoff}} = 7$ . Prior to computing ILC weights, all

maps are deconvolved from their beam (effectively blowing up noise on small scales, in particular for the lowest frequency channels). Then, the channels are combined using map modes for  $\ell < 300$ .

With the above choices, the reconstructed map can not be good on small scales. As the authors notice themselves, using small scale modes results in minimisation of noise rather than foregrounds (and, obviously, rejecting the low-frequency observations, which are the noisiest on small scales after deconvolution from the beam). Better results could probably be obtained by estimating weight maps for different bands of  $\ell$ . In essence, this is what our needlet ILC method permits to achieve.

As a last comment, we note that limiting the number of modes of the weight maps to  $\ell \leq 7$ , reportedly for reasons of singularity of the system to be solved, results in spatial coherence of the weights on scales of about 35 degrees. The galactic ridge, however, is about 1 degree thick. Hence, the spatial variability of the ILC weights achieved by Kim et al. [94] is not quite adapted to the actual scale of foreground variation. The needlet ILC method presented in our paper, as will be seen later on, solves this issue in a very natural way.

### Other maps

Other authors have performed various foreground cleaning in the WMAP observations, producing a number of CMB maps for several different models of the foreground emission. Bonaldi et al. [57] perform component separation on WMAP data using the CCA method described in Bonaldi et al. [56]. Maino et al. [103] obtain also several CMB maps, using the FastICA method [104]. None of the CMB maps obtained is full sky, nor publicly available yet. They are not considered further in this analysis.

Finally, some foreground cleaning has also been performed by Saha et al. [128]. Their paper also includes an interesting analysis of the ILC bias. The primary goal of that work, however, is to compute the CMB power spectrum, rather than producing a CMB map.

### Existing map summary

Figure 8.2 shows six available maps, all displayed in the same colour scale. It illustrates the resolution and foreground contamination of the various maps. Table 8.1 summarises the main properties of the maps. Only the TILC1 and TILC3 maps are high resolution attempts at component separation everywhere, including the galactic plane, combining all WMAP observations. All other maps (WILC maps, EILC1, EGS3, PILC3, KILC5) are at reduced resolution.

### 8.3.2 Map comparison

CMB maps produced from WMAP one year data have been compared by Eriksen et al. [75], showing quite large differences, ranging from  $-100$  to  $100 \mu\text{K}$ . Similarly, Park et al. [118] compare their PILC3 map with the WILC3 and TILC3 maps at 1.4 degree resolution, showing differences in excess of  $40 \mu\text{K}$ . In the following, discrepancies between these various solutions are further investigated.

As a first step, we evaluate by how much the various maps at one degree resolution disagree on the full sky. The top panel of Figure 8.3 shows the pixel based standard deviation of five such maps. In this evaluation, we exclude the EGS3 map (which is at 3 degree resolution, and is the source of additional

variance in the Kp2 sky mask, where the CMB is estimated at even lower resolution).

All maps are obtained from observation of the same sky (sometimes starting from the exact same data set) and are smoothed to the same resolution (one degree) where noise is small. Discrepancies originate essentially from systematic differences in the methods. In the central regions of the galactic plane, discrepancies are significantly higher than  $50 \mu\text{K}$  (with bright spots above  $90 \mu\text{K}$ ). In other sky regions, they are typically in the  $10\text{--}20 \mu\text{K}$  range, except on compact spots again, where they are above  $50 \mu\text{K}$ . The latter are probably due to residuals of the emission from strong compact sources. The discrepancy between the maps is larger in the ecliptic plane, which is a signature of the impact of the instrumental noise, uneven because of the WMAP scanning strategy. This is due to the different attention given to minimising the contribution of instrumental noise on small scales (rather than foregrounds) in the final map (and to a much lesser extent to the difference of noise level between the various WMAP releases).

The bottom panel of Figure 8.3 focuses the comparison to larger scales (3 degree beam). At high galactic latitudes, the discrepancies are below  $10 \mu\text{K}$  except for a few localised regions (LMC, Ophiuchus complex) where they reach about  $30 \mu\text{K}$ . Differences in the North Polar Spur, at the level of  $10\text{--}15 \mu\text{K}$ , are also clearly visible. Close to and inside the galactic plane, systematic discrepancies significantly exceed  $30 \mu\text{K}$ .

Including the EGS3 map for comparisons at high galactic latitudes does not change these conclusions. Pairwise comparisons of the available maps typically show the same level of discrepancies, which indicates that the variance of the solutions is not due to one single map in strong disagreement with all the others.

The conclusion of this comparison is that foreground residuals exist in the published CMB maps at the level of about  $50$  to  $100 \mu\text{K}$  in the galactic ridge,  $20$  to  $50 \mu\text{K}$  at low galactic latitudes, and  $10$  to  $20 \mu\text{K}$  at higher galactic latitudes (above  $30$  degrees).

This observation calls for localised weightings, adapted to local properties of the foregrounds and the noise. This, however, is not easily compatible with the recovery of the largest modes of the CMB map, as pointed out before by Eriksen et al. [75], and as demonstrated by the discontinuities observed between the CMB solutions in the different regions when the sky is cut, as in the TILC, EILC1, and WILC maps.

Our approach, then, will be to vary the relative weightings on small scale for small scale CMB reconstruction, and keep the weighting uniform over large regions of the sky for the recovery of the largest scales. This can be achieved quite straightforwardly by using the spherical needlets discussed in Section 8.4.

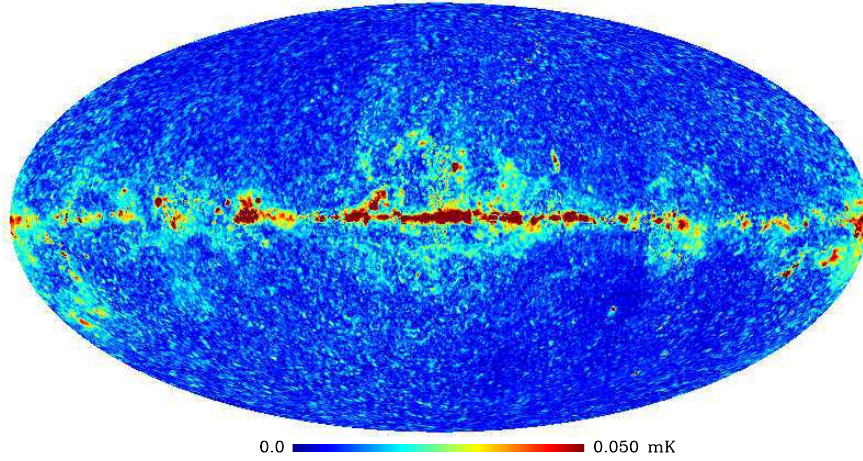
## 8.4 The ILC needlet method

### 8.4.1 The choice of the ILC

It is striking that all the presently available full sky CMB maps derived from an analysis of the WMAP data have been obtained by an implementation of the ILC method.

The ILC, indeed, has many advantages :

Standard Deviation of EILC1, TILC3, PILC3, WILC5 and KILC5 (1 deg. resolution)



Standard Deviation of EILC1, TILC3, PILC3, WILC5 and KILC5 (3 deg. resolution)

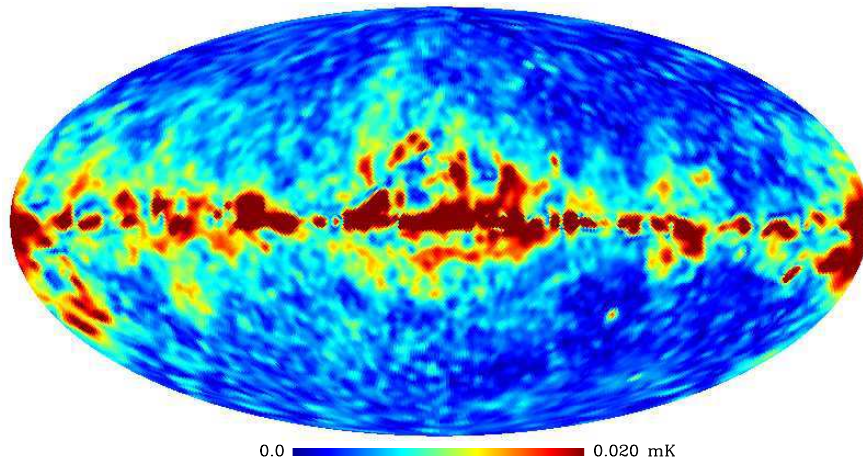


FIG. 8.3 – Top : the standard deviation, per pixel for  $n_{\text{side}}=512$ , of the EILC1, TILC3, PILC3, WILC5 and KILC5 maps, at  $1^\circ$  resolution. Bottom : same at  $3^\circ$  resolution. Note the different color scales for the two panels.

- The method relies only on two very safe assumptions : the CMB emission law, and the fact that the CMB template is not correlated to foreground emission<sup>3</sup> ;
- Under these assumptions, the method minimises the empirical variance of the reconstruction error ;
- The ILC is very simple in implementation.

The ILC has also two major drawbacks.

- As noted for instance by Hinshaw et al. [86], Delabrouille and Cardoso [67], Saha et al. [128], empirical correlations between the CMB and the source of contamination results in a bias ; this bias is discussed in more detail below ;
- In absence of a model of the contaminants (foregrounds and noise), it is not possible to predict the reconstruction errors, which somewhat annihilates the benefit of making very safe assumptions about the properties of the data set.

### 8.4.2 The ILC bias

The existence of a “bias” in maps obtained by an ILC method is a well established fact. The derivation of this bias (which is, in fact, the systematic cancelling of a fraction of the projection of the CMB map onto the vector space spanned by the noise realisations for all the considered input maps), is given in appendix.

The order of magnitude to keep in mind is that about  $(m - 1)$  “modes” of the original CMB, out of  $N_p$ , are cancelled by the ILC, where  $m$  is the number of channels used, and  $N_p$  the number of independent pixels or modes in the regions for which the ILC is implemented independently. Note however that when the signal is strongly correlated between pixels, the bias can be significantly larger –see the appendix for details.

The practical consequences are :

- a loss of CMB power, which has to be taken into account for power spectrum estimation ;
- an anti-correlation of the map reconstruction error with the real CMB sky.

The level of the bias induced by our method is investigated both theoretically (Appendix 8.9), and through Monte-Carlo simulations.

### 8.4.3 Needlets

A *frame* is a collection of functions with properties close to those of a basis. *Tight frames* share many properties with orthonormal bases, but are redundant (see [9] for details).

Needlets were introduced by Narcowich et al. [33] as a particular construction of a wavelet frame on the sphere. They have been studied in a statistical context (e.g. Baldi et al. [151, 149]) and have also been used recently for cosmological data analysis problems (e.g. Pietrobon et al. [121]). The most distinctive property of the needlets is their simultaneous perfect localization in the spherical harmonic domain (actually they are spherical polynomials) and potentially excellent localization in the spatial domain.

We recall here the definition and practical implementation of the needlet coefficients, following the generalised formulation by Guillaux et al. [25]. Let  $b_{j,\ell}$ ,  $j \in \mathcal{J}$  be a collection of window functions in

---

<sup>3</sup> In reality, it is likely that the CMB map actually *is* somewhat correlated to the foregrounds (extragalactic point sources and SZ effect), because of the ISW effect. The implication of this is not studied further in the present paper.

the multipole domain, indexed by  $j$ . Suppose that for each scale  $j$ ,  $\xi_{j,k}$  is a grid of points (indexed by  $k \in K^{(j)}$ ) which satisfies an exact<sup>4</sup> quadrature condition with weights  $\lambda_{j,k}$ . The needlets are axisymmetric functions defined by

$$\psi_{j,k}(\xi) = \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\ell^{max}} b_{j,\ell} L_{\ell}(\xi \cdot \xi_{j,k}), \quad (8.6)$$

where  $L_{\ell}$  denote the Legendre polynomials.

Any square integrable function  $f$  on the sphere can be analysed by the scalar products  $\beta_{j,k} := \langle f, \psi_{j,k} \rangle$  of the function  $f$  with *analysis needlets*. All the needlet coefficients of scale  $j$  are advantageously computed in the spherical harmonic domain, as the evaluation at points  $\xi_{j,k}$  of a function whose multipole moment are simply  $b_{j,\ell} a_{\ell m}$ . These needlet coefficients, denoted  $\gamma_{j,k}$ , are given by :

$$\gamma_{j,k} = (\lambda_{j,k})^{-1/2} \beta_{j,k}$$

Each field of needlet coefficients can in turn be convolved with some *synthesis needlets*

$$\tilde{\psi}_{j,k}(\xi) = \sqrt{\lambda_{j,k}} \sum_{\ell=0}^{\ell^{max}} \tilde{b}_{j,\ell} L_{\ell}(\xi \cdot \xi_{j,k}), \quad (8.7)$$

using the exact same procedure and leading to the map  $X^{(j)}$  whose multipole moments are  $b_{j,\ell} \tilde{b}_{j,\ell} a_{\ell m}$ . The analysis and synthesis operations are summed up as :

*Analysis :*

$$X \xrightarrow{\text{SHT}} a_{\ell,m} \xrightarrow{\times} b_{j,\ell} a_{\ell,m} \xRightarrow{\text{SHT}^{-1}} \gamma_{j,k}$$

*Synthesis :*

$$\gamma_{j,k} \xRightarrow{\text{SHT}^{-1}} b_{j,\ell} a_{\ell,m} \xrightarrow{\times} \tilde{b}_{j,\ell} b_{j,\ell} a_{\ell,m} \xRightarrow{\text{SHT}} X^{(j)}$$

Double arrows denote as many transforms as scales in  $\mathcal{J}$ . If  $X$  is band-limited to  $\ell \leq \ell^{max}$  and if the reconstruction condition  $\sum_j b_{j,\ell} \tilde{b}_{j,\ell} = 1$  holds for all  $\ell \leq \ell^{max}$ , then the complete process yields a decomposition of  $X$  in smooth maps, namely

$$\forall \xi, \quad X(\xi) = \sum_j X^{(j)}(\xi) \quad (8.8)$$

Note that the existence of a fast inverse spherical harmonic transform using the quadrature points  $\xi_{j,k}$  is required in practice, and that HEALPix pixels and weights fulfil the quadrature condition only approximately. Further details can be found in Guillaux *et al.* [25], with an extensive discussion on the choice of the spectral window functions.

A key feature of the needlet decomposition follows from the localization of the analysis functions which allows for localised processing (such as denoising, signal enhancement, masking etc.) in the needlet coefficient domain, *i.e.* applying some non-uniform transforms to the coefficients  $\gamma_{j,k}$ . Other types of wavelets (for instance, the steerable wavelets of Wiaux *et al.* [145], or the wavelets, ridgelets and curvelets of Starck *et al.* [134]) could also be used for localised processing on the sphere,

---

<sup>4</sup>or almost exact, for all practical purposes



although the quality of the localisation depends much on the details of the wavelet design. Needlets are compactly supported in the multipole domain and can be further designed to be well localised in direct domain according to various criteria. This permits to work on full sky data without real need for masking the galactic ridge.

#### 8.4.4 The method

The method implemented in this work, and applied both to simulations and to the real WMAP data sets (for all releases), consists in the following steps :

- We start with the data set consisting of band-averaged temperature maps from WMAP (simulated or real data), to which we add the IRIS 100 micron map ;
- WMAP-detected point sources are subtracted from the WMAP maps ;
- We apply a preprocessing mask, in which a very small number of very bright, compact regions, are blanked (see Table 8.1) ; blanked regions are filled-in by interpolation ; this is done only on the real WMAP data.
- All maps are deconvolved to the same resolution (that of the W channel<sup>5</sup>) ; this operation is performed in harmonic space ;
- Maps are analysed into a set of needlet coefficients  $\gamma_{j,k}$  following the method described in 8.4.3 ;
- For each scale, the covariance matrix  $\widehat{R}$  of the observations is computed locally (using an average of  $32 \times 32$  needlet coefficients) ;
- The ILC solution is implemented for each scale in local patches ;
- An output CMB map is reconstructed from the ILC filtered needlet coefficients ; this map constitutes our main CMB product ;
- That map is Wiener-filtered in harmonic space, to make an alternate CMB map with lower integrated error (our best guess CMB map).
- In parallel, the actual ILC filter used on the analysed data set is applied to 100 different simulations of the WMAP noise, to estimate the noise contribution to the final map.
- The level of the biasing, which depends on the geometry and not much on the actual templates of CMB and foreground and noise, is estimated on a set of fully simulated data.

Each of these steps is described in more detail below.

#### 8.4.5 Point source subtraction

Strong point sources in the input data set typically leave detectable residuals in the output ILC map, and twist at the same time the estimation of the background, thus conducting to lower rejection of other contaminants. On the other hand, their specific shape usually allows effective estimation and removal by other methods. For the purpose of this study we used information from the WMAP source catalogue (Hinshaw et al. [86]) which provides characterisation for all point sources detected above a  $5\sigma$  threshold away from the galactic plane.

For all sources identified, we subtract from the input maps a Gaussian profile at the given position and with the given flux. Conversion factors between flux density and Gaussian amplitude, as well as the FWHM of the Gaussian profile are taken from Table 5 of Page et al. [117] for one year data, and corresponding updates for the more recent releases.

---

<sup>5</sup>A noise weighted average beam is obtained from the W1 W2 W3 W4 beam coefficients provided by the WMAP team

Name	Galactic coordinates	Type
Crab neb	184.5575 -05.7843	SNR
sgr A	000.064 +00.147	Radio-Source
sgr B	000.599 +00.002	Radio-Source
sgr C	359.4288 -00.0898	HII region
sgr D	001.131 -00.106	Molecular cloud
Orion A	209.0137 -19.3816	HII region
Orion B	206.5345 -16.3539	Molecular cloud
Omega neb	015.051 -00.674	HII region
Cen A	309.5159 +19.4173	QSO
Cas A	111.735 -02.130	SNR
Carina neb	287.6099 -00.8542	HII region

Table 8.1 – List of compact regions blanked in the pre-processing step. A circular patch centered on the source, of radius 75, 55, 45, 45, 34 and 35 arcminute for the K, Ka, Q, V, W and IRIS 100  $\mu m$  bands respectively, is masked. The masked regions are then filled with an extrapolation of edge values.

For the simulated data set, the subtraction of detected point sources is mimicked by removing from the simulation all sources above 1 Jy (independently in all channels).

#### 8.4.6 Blanking of compact regions

In addition to the point sources subtracted above, some compact regions of strong emission (mostly in the galactic plane) exceed the rejection capabilities of the method used in this analysis, because they are too local and/or too specific. Their contribution in the wings of the needlets also contaminate the solution far from the centre of the sources. Those sources cannot be satisfactorily subtracted in the same way than the previous ones, either because they are not strictly speaking point-like, or because they are bright enough that small departures of actual beam shapes from the Gaussian model used in the subtraction step leave significant residuals. As they represent only a very tiny fraction of the sky (we single out eleven such sources), we blank out these regions in all WMAP channels, cutting out circular patches adapted to the size of the beam and of the source. Table 8.1 gives the list of those regions with their main characteristics.

To reduce local pollution of the needlet coefficients by the sharp cut, the small blanked regions are filled in by a smooth interpolation, so that fluctuations at a larger scale than the hole size are at least coarsely reconstructed. More precisely, interpolation is made by diffusion of the boundary values inside the hole.

Although this masking and interpolation has no reason to be optimal, it is an efficient way of reducing the impact of very strong sources on their environment. The CMB inside the masked patches is recovered (to some extent) both by the interpolation of original maps (which avoids sharp discontinuities) and by the needlet decomposition and ILC reconstruction. The masked region is tiny : 0.058% of the sky in the K channel (the most affected).

Band ( $j$ )	$\ell_{\min}$	$\ell_{\max}$	$\text{nside}(j)$
1	0	15	8
2	9	31	16
3	17	63	32
4	33	127	64
5	65	255	128
6	129	511	256
7	257	767	512
8	513	1023	512
9	769	1199	512

Table 8.2 – Spectral bands used for the needlet decomposition in this analysis. Needlet coefficient maps are made at different values of  $\text{nside}$ , given in the last column.

#### 8.4.7 Needlet decomposition

The original observations (WMAP and IRIS) are decomposed into a set of filtered maps represented by their spherical harmonic coefficients :

$$a_{\ell m}^{(j)} = b_{j,\ell} a_{\ell m} \quad (8.9)$$

where  $a_{\ell m}$  are the spherical harmonics coefficients of the original map, and  $a_{\ell m}^{(j)}$  those of the same map filtered by the window function  $j$ . Needlet coefficients  $\gamma_{j,k}$  are obtained as the value of the filtered map at points  $\xi_{j,k}$ .

For each scale  $j$ , the coefficients  $\gamma_{j,k}$  are computed on a HEALPix grid at some value of  $\text{nside}$ , compatible with the maximum value of  $\ell$  of band  $j$ . We use for  $\text{nside}(j)$  the smallest power of 2 larger than  $\ell^{\max}/2$ , with a maximum of 512. Details about the bands used are given in table 8.2 and figure 8.4.

#### 8.4.8 ILC implementation on needlet coefficients

The general idea is to implement independently the ILC on subsets of the needlet coefficients  $\gamma_{j,k}$ . For a given scale, these coefficients come in the format of a set of HEALPix maps (one per frequency channel). The ILC is implemented locally in space and locally in  $\ell$  as follows.

Covariance matrices  $\mathbf{R}_k^{(j)} = \langle \gamma_k^{(j)} \gamma_k^{(j)T} \rangle$  for scale  $j$  at pixel  $k$  are estimated as the average of the product of the computed needlet coefficients over some space domain  $\mathcal{D}_k$ . Because of this, there is a trade-off between localization and accuracy of estimation. A better estimate of the true covariance is obtained by averaging the products of needlet coefficients between all pairs of maps over a large area, but this provides an estimate of the average covariance over that area. In practice, for the present analysis, we make use of the hierarchical property of the HEALPix pixellisation, and compute covariance matrices as the average in larger pixels, corresponding to a HEALPix pixellisation with  $\text{nside} = \text{nside}(j)/32$ . This provides a computation of the statistics by averaging  $32^2 = 1024$  samples, which results in a precision of order 3 per cent for all entries of the  $\mathbf{R}_k^{(j)}$  matrix. It implies an ILC bias of order  $5/1024$  (for  $m = 6$  channels and  $N_p = 1024$  coefficients per domain on which the ILC filter is estimated independently, see appendix for details). Choosing a smaller area results in excessive error in the

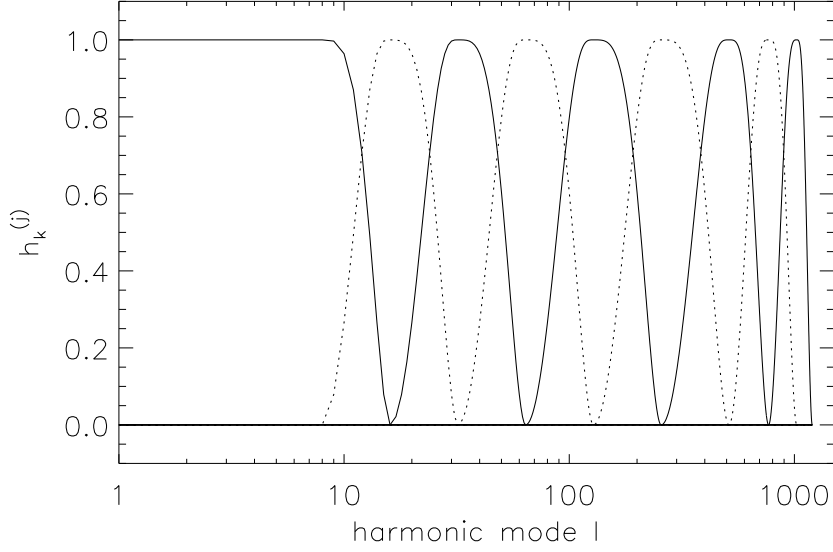


FIG. 8.4 – The spectral bands used in this work for the definition of the needlets.

covariance estimates, and hence excessive bias. Choosing a larger area results in less localisation, and hence some loss of efficiency of the needlet approach.

We denote as  $\widehat{\mathbf{R}}_{k \in \mathcal{D}_k}^{(j)}$  the estimate of  $\mathbf{R}_{j,k}$  obtained by averaging the value of  $\boldsymbol{\gamma}_{j,k'} \boldsymbol{\gamma}_{j,k'}^T$  in domain  $\mathcal{D}_k$ .

On the largest scales ( $\ell \leq 50$ ), the typical angular extent of a needlet is larger than 5 degrees, and the value of `nside` for the map of needlet coefficients is less than 32. Covariance matrices are then computed on the full sky rather than on the largest possible HEALPix grid, i.e.  $\mathcal{D}_k$  is the complete sky, rather than one of the 12 basis Healpix pixels.

Using these covariance matrices, the ILC is implemented using equation 8.4 for each domain. The estimated CMB needlet coefficients are :

$$\left[ \widehat{\boldsymbol{\gamma}}_{j,k} \right]_{\text{CMB}} = \frac{\mathbf{a}^T \left[ \widehat{\mathbf{R}}_{k' \in \mathcal{D}_k}^{(j)} \right]^{-1}}{\mathbf{a}^T \left[ \widehat{\mathbf{R}}_{k' \in \mathcal{D}_k}^{(j)} \right]^{-1} \mathbf{a}} \boldsymbol{\gamma}_{j,k} \quad (8.10)$$

#### 8.4.9 Full map reconstruction

The full CMB map reconstructed from this set of needlet coefficients is our basic needlet ILC (NILC) CMB map.

#### 8.4.10 Final Wiener filtering

For a number of purposes, in particular subtraction of an estimate of the CMB to study other emissions, it is interesting to use, instead of our ILC map at the resolution of the WMAP W channel, a map with minimal error. Such a map is obtained from the ILC map by one-dimensional Wiener filtering.

As a last processing step towards a minimum variance CMB map, we thus Wiener-filter our CMB map, to get rid of the large noise contamination at high  $\ell$ . The Wiener filter is performed in harmonic space as described in 8.2.3.

The harmonic Wiener filter is given by formula 8.5, i.e.  $w_\ell = B_\ell^2 C_\ell / (B_\ell^2 C_\ell + N_\ell)$ . For its implementation, we need to know the relative power of CMB and noise. We assume that the best fit CMB power spectrum of the WMAP team is correct, hence  $C_\ell$  is known. The beam factor  $B_\ell$  is assumed perfectly known as well. The denominator  $B_\ell^2 C_\ell + N_\ell$  can be estimated directly as the power spectrum of our output needlet ILC map.

In practice, we smooth the power spectra with  $\delta\ell/\ell = 0.1$  to lower the variance of the power spectrum estimator on the output needlet ILC map. Even with this, the filter is poorly estimated for low modes, because of the large cosmic variance. As can be seen on Figure 8.6, the signal to noise ratio of our reconstructed map is expected to be quite high at low  $\ell$ . Therefore, the Wiener filter for low modes is expected to be very close to 1. For this reason, we set  $w_\ell = 1$  for  $\ell < 200$ , and use a linear interpolation between  $\ell = 200$  and  $\ell = 250$ .

#### 8.4.11 Noise level estimate

The level of noise contamination (variance per pixel, and average power spectrum) in the output map is estimated by Monte-Carlo, using the average of 100 realisations of the WMAP noise maps. For each initial set ( $i$ ) of five noise maps (one noise map per WMAP channel), a single output noise map  $n_p^{(i)}$  is obtained by performing the needlet decomposition of the initial noise maps, and filtering needlet coefficient maps with the same filter as that used on the single full simulated data set.

Denoting as  $n_p^{(i)}$  and  $n_{\ell m}^{(i)}$  respectively the pixel value and the harmonic space value of the noise map number  $i$ , we compute :

$$\sigma_p^2 = \frac{1}{N_i} \sum_i (n_p^{(i)})^2 \quad (8.11)$$

and

$$\sigma_\ell^2 = \frac{1}{N_i(2\ell + 1)} \sum_i \sum_m (n_{\ell m}^{(i)})^2 \quad (8.12)$$

These are respectively estimates of the noise pixel variance and of the noise power spectrum of our final map.

#### 8.4.12 Bias estimates

The impact of the ILC bias is estimated by Monte-Carlo simulations on simulated data sets. The corresponding error is of order 2% of the CMB.

#### 8.4.13 Practical implementation

The practical implementation of this processing pipeline is made essentially using the octave language (the free software version of Matlab). The analysis is done in the framework of the pipeline tool

developed in the ADAMIS team at the APC laboratory. This tool provides a flexible and convenient web interface for running our data analysis on simulations or real data with easy handling and tracing of the various pipeline options<sup>6</sup>. Single runs of the full pipeline require less than half an hour on a single processor of a standard desktop computer (dominated by harmonic transforms), whereas numerous pipelines on simulated data sets for Monte-Carlo are run on the ADAMIS 88-processor cluster, optimised for efficient I/O.

## 8.5 Simulations

### 8.5.1 The simulated data

We start with a validation of our method on simulated data sets. For this experiment, synthetic observations of the sky emission are generated using the Planck Sky Model (PSM). The PSM is a flexible software library, designed for simulating the total sky emission in the 10-1000 GHz frequency range, and developed as part of the foreground modelling activities of the Planck working group on component separation (Planck WG2). Sky emission comprises galactic components of four origin (free-free, synchrotron, thermal dust, and spinning dust, with spectral emission laws for dust and synchrotron varying from pixel to pixel), CMB, kinetic and thermal SZ effect, and the emission from a population of galactic and extragalactic point sources which includes radio sources, infrared sources, and an infrared background. Although not perfect, the model sky is thought to be sufficiently representative of the complexity of the real sky emission for our simulations to be meaningful.

Sky maps are produced at WMAP central frequencies for the K, Ka, Q, V and W band, and convolved in harmonic space with approximate WMAP instrumental beams (Gaussian symmetric beams are used for these simulations). Uncorrelated, non-stationary Gaussian noise is added, with a pixel variance deduced from the WMAP sensitivity per channel and effective hit count. To mimic the subtraction of the brightest point sources detected by WMAP, we remove from the model sky, at each frequency, all sources with flux above 1 Jy (assuming they would have been detected, and can be subtracted from the data set). The 11 compact regions listed in table 8.1 however, being specific to the real sky, are not blanked for the simulations.

Although these simulations provide only an approximation of the real WMAP data sets, they are representative enough that the simulated data offer a component separation challenge close to that of the real data set. The IRIS map is used as part of the full set of data for the ILC implementation on simulations.

### 8.5.2 Results

Figure 8.5 shows the input simulated CMB, the output CMB, and the difference of the two for one particular simulation. The reconstruction is visually good except in regions of local strong galactic emission (in the galactic ridge, for example). This is to be expected : not only the method can not remove foregrounds perfectly, but in addition the price to pay to remove foregrounds (even imperfectly) is more noise (because of sub-optimal weighting of the observations as far as noise contamination is concerned).

---

<sup>6</sup>see [http://www.apc.univ-paris7.fr/APC\\_CS/Recherche/Adamis/](http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/) in the ‘outreach’ section

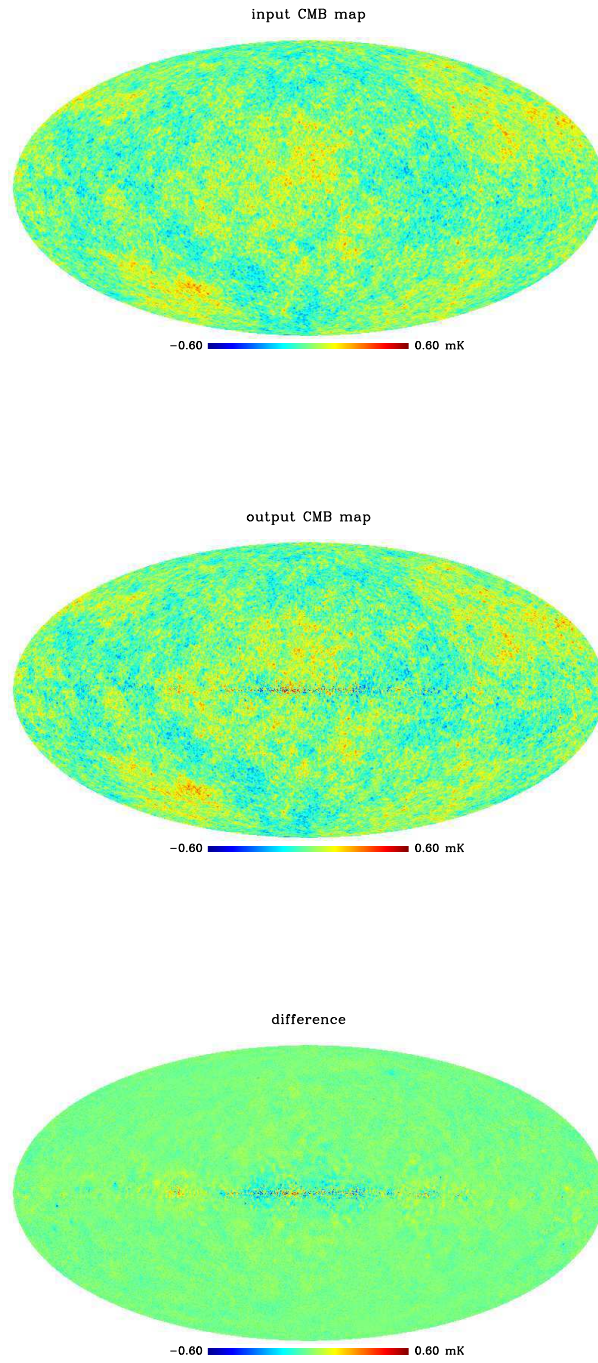


FIG. 8.5 – Top : the simulated input CMB map. Middle : the reconstructed CMB. Bottom : the difference (output-input), displaying the residuals left by the method. All three maps share the same colour scale, and are at the resolution of the WMAP W channel.

A more quantitative estimate of the level of contamination of the CMB map by foregrounds and noise is obtained by looking at power spectra. Figure 8.6 shows the input simulated CMB power spectrum (dotted line), the spectrum of the output CMB (solid black line), and the spectrum of the map of residuals (difference between output and input, dashed line), both full sky (top panel) and in the HGL region (bottom panel). The angular power spectrum of the residual map is seen to be small as compared to the CMB power on large scales, the two being comparable at  $\ell \simeq 500$ . Noise dominates on smaller scales. The residuals due to the presence of galactic emission are seen to contribute power essentially below  $\ell = 400$ , where the power of the difference map is seen to be slightly higher in the full sky power spectrum than in the HGL power spectrum (this is visible, in particular, at the top of the first acoustic peak).

### 8.5.3 Bias

As discussed in 8.4.2, we expect a (small) bias in the ILC map, due to empirical correlations between the CMB emission and contaminants (including noise and foregrounds). This is not particular to our approach, and is expected for any ILC method. For better characterisation of our output map, we evaluate the effect both theoretically (in appendix) and numerically.

Although a general analytic estimate of the bias is complicated, appendix 8.9 shows that (to first order at least and for reasonable assumptions about the CMB, foregrounds, and noise) the amplitude of the effect does not depend much on what the actual foregrounds are in detail, but is set essentially by the geometry of the domains considered (through a number of effective modes). It is then possible to estimate the amplitude of the effect by Monte-Carlo simulations on synthetic data sets resembling the actual WMAP observations.

Figure 8.7 illustrates an estimate of the bias  $b(\ell)$  as a function of the harmonic mode, computed as a fractional error :

$$b(\ell) = \frac{\sum_m (\widehat{a}_{\ell m} - a_{\ell m}) a_{\ell m}^*}{\sum_m |a_{\ell m}|^2} \quad (8.13)$$

where  $a_{\ell m}$  are the harmonic modes of the input CMB map, and  $\widehat{a}_{\ell m}$  the harmonic modes of the output CMB map. The numerator in this equation computes the covariance of the residual map and the input map as a function of  $\ell$ , and the denominator is a normalisation factor. For an error uncorrelated with the input,  $b(\ell)$  should be close to 0 on average. Analytical estimates of the effect (see appendix) suggest a bias of order 2% for our implementation (taking into account mode correlation). The numerical estimate of Figure 8.7, obtained as the average bias for 500 simulated data sets, is in good agreement with this prediction, with slight variations due to varying number of effective modes for different needlet scales. The average value of  $b(\ell)$  between  $\ell = 2$  and 1000 on that simulation is about  $-2.2\%$ .

## 8.6 Application to WMAP data

### 8.6.1 ILC Result

We now turn to the description of the results obtained on the real WMAP data sets. In order to facilitate the comparison with existing maps, we process independently one year, three year and five year data,



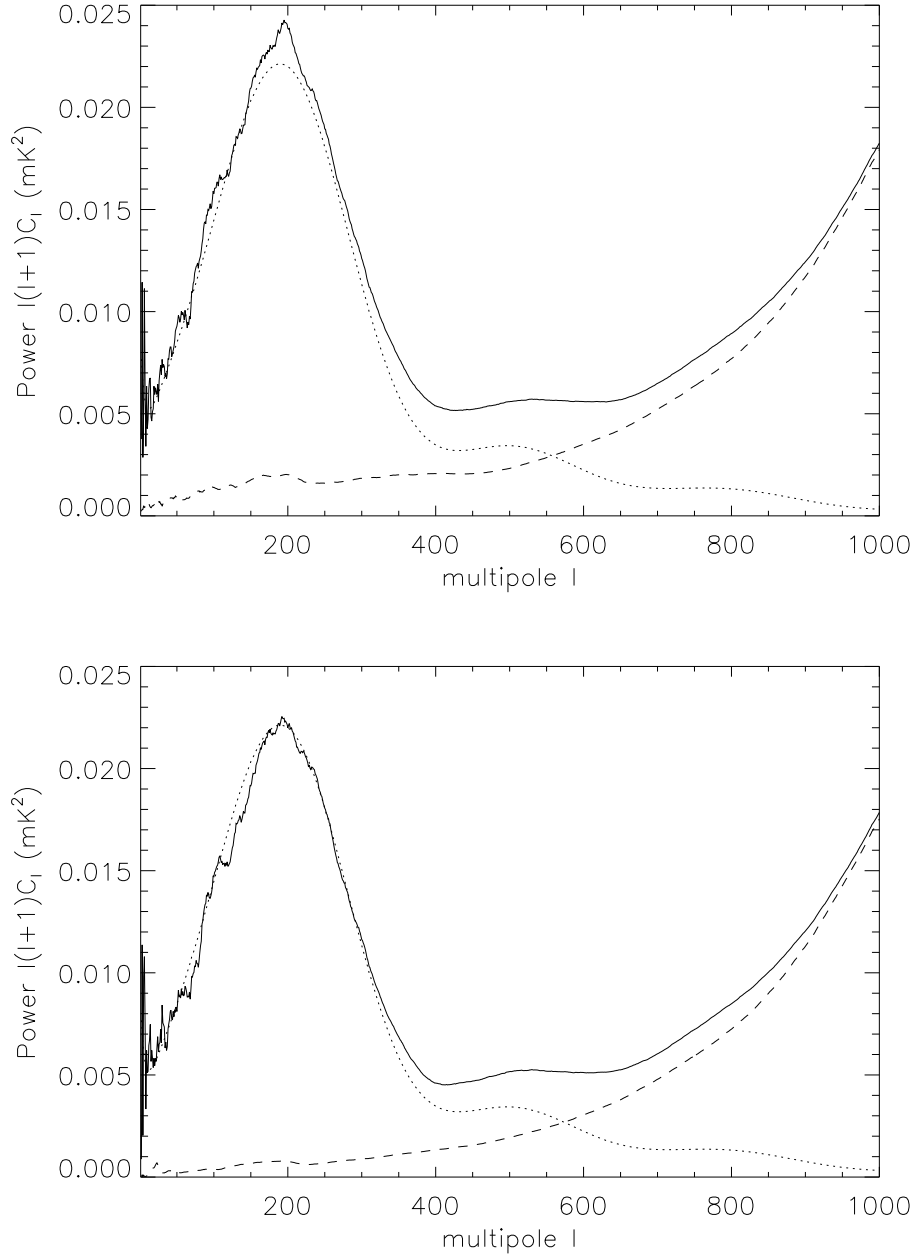


FIG. 8.6 – Top : For simulated data sets, full sky power spectra of output CMB map (plain line), and of difference map (dashed line). The CMB model used in the simulation is over-plotted as a dotted line. Bottom : Same at high galactic latitude only (HGL region).

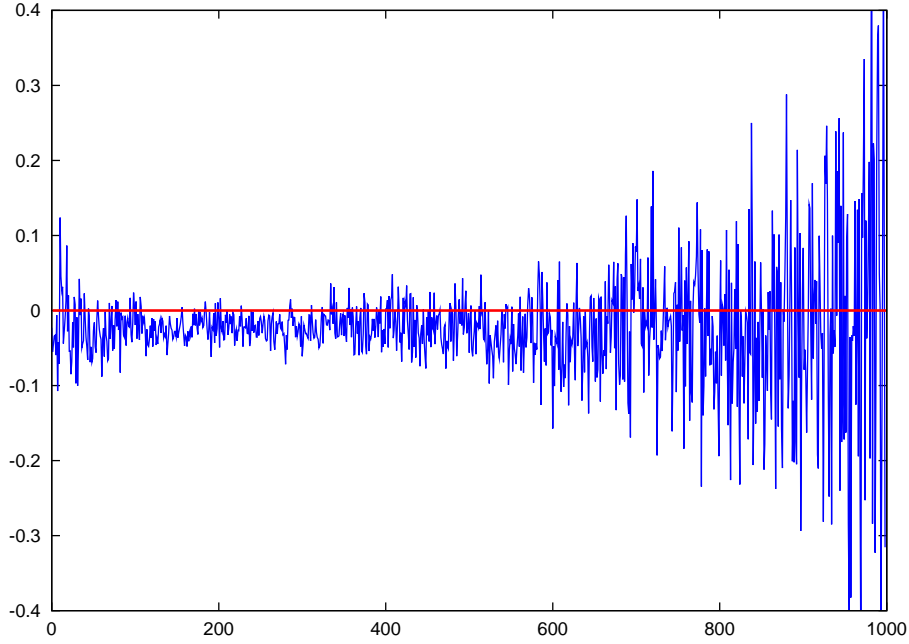


FIG. 8.7 – The fractional bias in the ILC map as a function of  $\ell$ , for one single simulation. This figure is illustrative both of the amplitude of the effect, and on its variance for one single realisation. Bias, and standard deviation of the bias, are of the same order of magnitude for most of the  $\ell$  range.

to obtain three CMB maps (hereafter NILC1, NILC3 and NILC5). For each year, we use the beam estimates, noise level, and point source catalogue provided with the corresponding release.

The improvement of CMB reconstruction with consecutive data releases is illustrated on Figure 8.8, which shows the full sky power spectra of the NILC CMB maps obtained with one year, three year, and five year WMAP data. The power spectra displayed are the raw power spectra of the output map, computed full sky, and smoothed with a variable window in  $\ell$  of 10% width. Whereas the lower part of the spectrum, cosmic variance limited and CMB dominated, does not change much, the high  $\ell$  spectrum of the map, dominated by noise, decreases substantially with increasing observation time – as expected. The excellent agreement at low  $\ell$  (up to  $\ell \simeq 300$ ) between the power spectra and the model is striking. The bumps in the spectrum due to the first and second acoustic peaks are clearly visible on the five year map spectrum.

Our full-sky reconstructed CMB map for the five year observations, at the resolution of the W-channel, is displayed in the top panel of Figure 8.9.

## 8.7 Discussion

### 8.7.1 Comparison with other maps

A full comparison of our needlet ILC maps (for all data releases) with all the other available maps would be too long for the present discussion. Rather, we decide to compare our five year result only

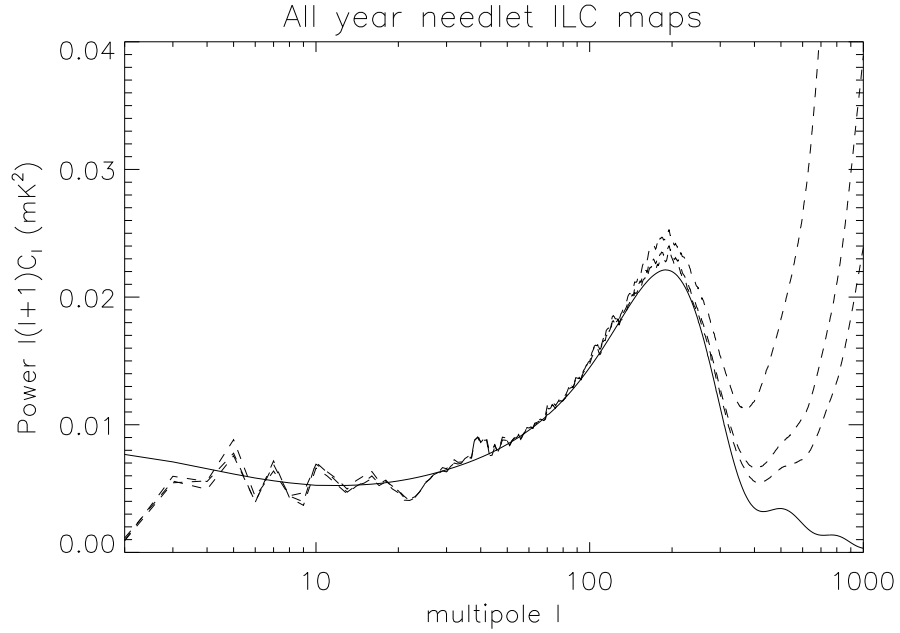


FIG. 8.8 – Power spectra of the NILC map for one year, three year, and five year WMAP data

with the TILC3 map on small scales (choice is motivated by the fact that the TILC is the only other full sky high resolution map available), and with the EGS3 map on large scales. This is of particular interest, as the EGS3 is the only map obtained with a method not based on the ILC, and also is a method specifically implemented for the recovery of the largest scales.

### Comparison at the pixel level – small scales

On the smallest scales, we compare our needlet ILC map with the TILC and with the WMAP foreground-reduced W band. Figure 8.11 shows local regions of the foreground-reduced map, the NILC5 map, and the TILC3 map, in the galactic plane and at the galactic pole. Our needlet map is clearly less contaminated by galactic emission than the other two. At high galactic latitude, the NILC5 and TILC3 are visually comparable, while the foreground-reduced map appears to be more noisy, as expected.

The power spectrum of the output map for the five year data (NILC5 map), for three different sky coverages, is shown on the bottom panel of Figure 8.10. On the same panel, we plot the angular power spectrum  $C_\ell$  corresponding to the WMAP best fit model, corrected for the W-channel beam. On the top panel of the same figure, we show the same power spectrum estimates for the TILC3 map. This shows the improvement achieved by our method close to the galactic plane. This improvement is due both to the needlet approach and to the use of the IRIS map to help with dust subtraction. As seen in Figure 8.8, the difference in quality between NILC5 and TILC3 cannot be explained solely by reduced noise (NILC5 and NILC3 being very close in quality for all scales except the smallest ones).

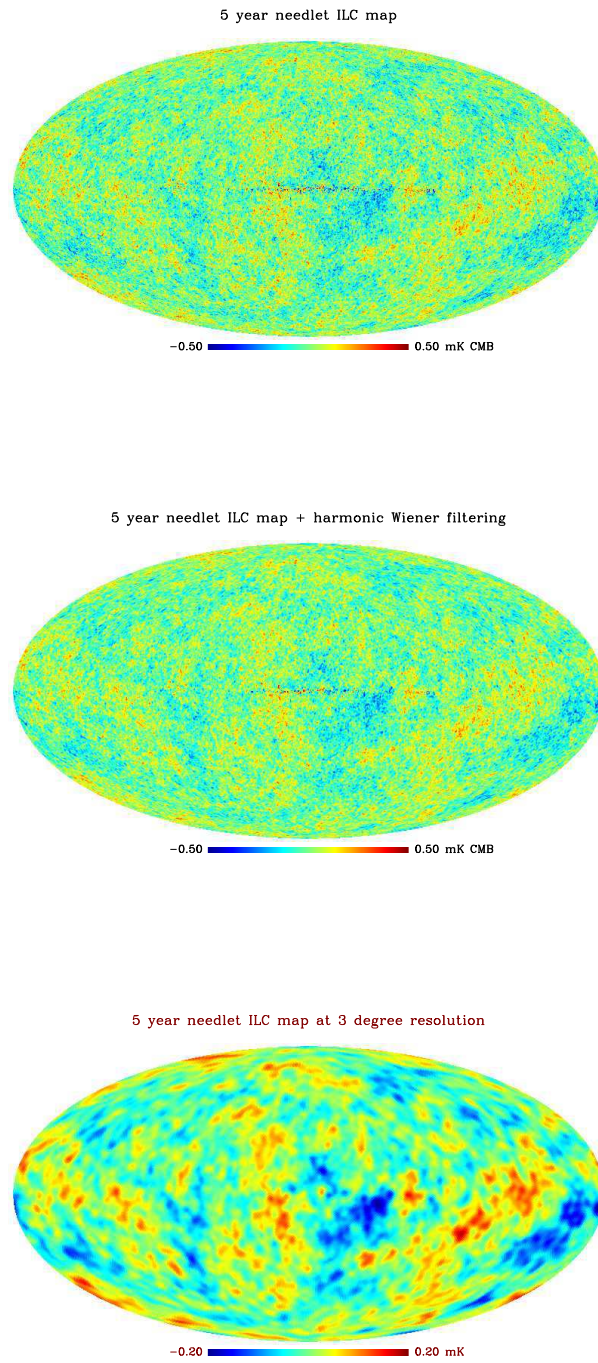


FIG. 8.9 – Top : the NILC5 reconstructed WMAP CMB at the resolution of the W channel. Middle : the harmonic Wiener NILC5 CMB map. Bottom : The NILC5 CMB map at 3 degree resolution.

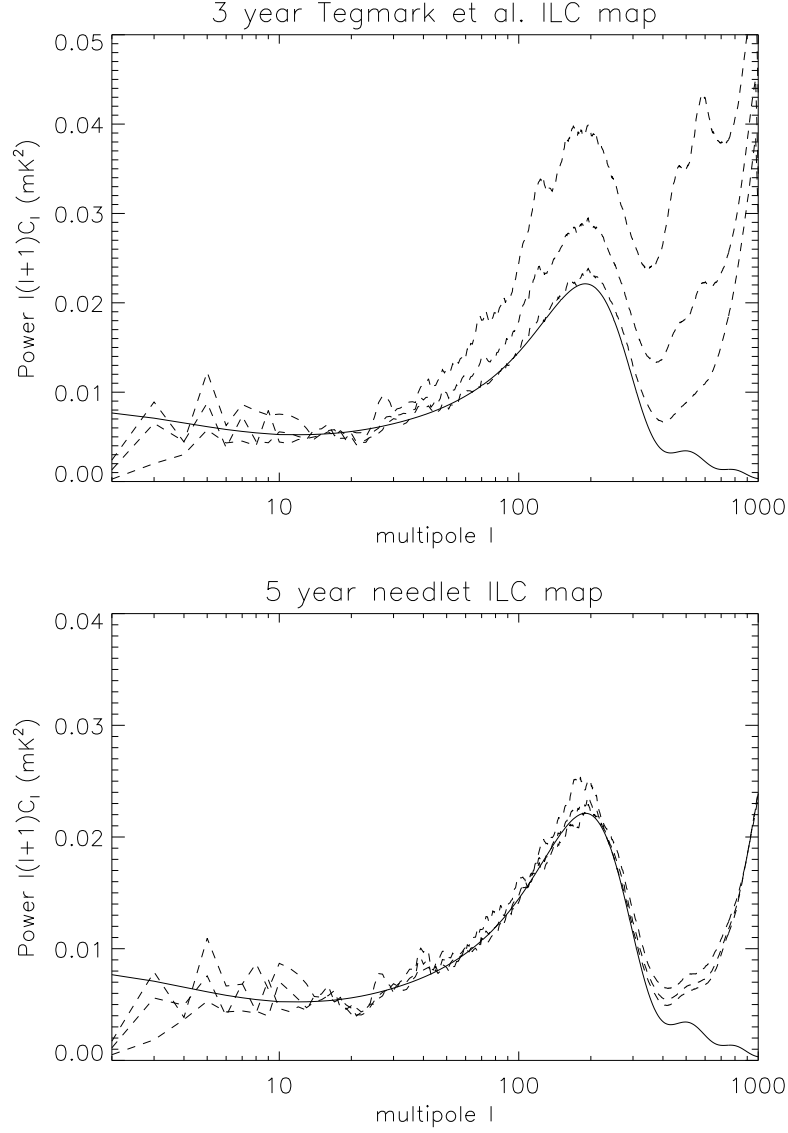


FIG. 8.10 – Power spectrum of the reconstructed WMAP CMB map. For each of the two panels, the CMB best fit model is shown as a solid black line, and power spectra computed at low galactic latitudes (using the LGL mask), on the full sky (no mask), and at high galactic latitudes only (HGL mask) are displayed as dashed lines. Note that the map spectra plotted here are directly those of the maps, without any correction for the beam. Top panel : ILC map of Tegmark et al. [139] (TILC3). Bottom panel : this work, with five year WMAP data. The power spectrum of the needlet ILC CMB is significantly more homogeneous than the power spectrum of the TILC3 map. We interpret this difference as an indication that the TILC3 map is significantly more contaminated by residuals of galactic emission. Note the different scales of the y-axis, and the improvement on small scales, with a noise power of about  $0.024 \text{ mK}^2$  at  $\ell = 1000$  for the NILC5 map at high galactic latitude, instead of about  $0.040 \text{ mK}^2$  for TILC3. As indicated by Figure 8.8, this is due essentially to the better quality of the five year release, since the NILC3 map also has a noise power spectrum of about  $0.040 \text{ mK}^2$  at  $\ell = 1000$ .

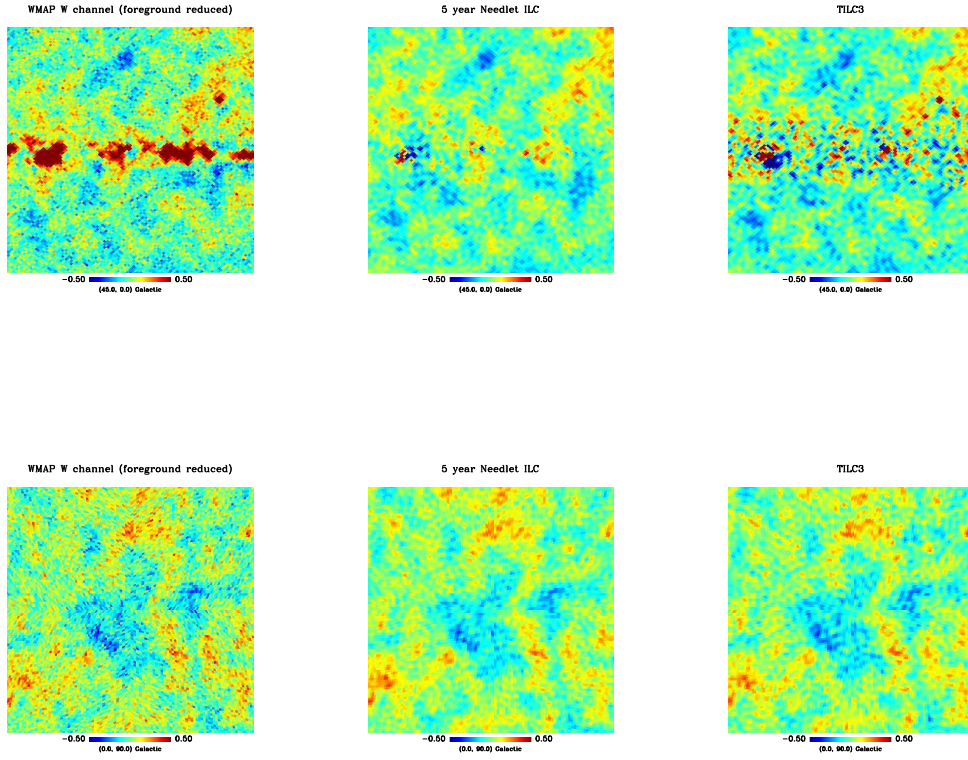


FIG. 8.11 – CMB maps from WMAP three year data obtained with three techniques : left column : WMAP foreground reduced, W channel ; middle column : our needlet ILC map ; right column : the TILC map. The top line corresponds to a patch located in the galactic plane, centred around coordinates  $(l, b) = (45^\circ, 0^\circ)$ . The bottom line shows the recovered CMB around the North Galactic Pole.

## Comparison at the pixel level – large scales

Figure 8.12 gives a visual comparison of NILC5 (this work) and the EGS3 [76], as well as of NILC5 and KILC5 [94]. On the top row, we display the EGS3 and KILC5 at a resolution of 3 degrees, and degraded to  $n_{\text{side}} = 64$ . The bottom row shows the difference between our needlet ILC solution (displayed on the bottom panel of Figure 8.9) and these two maps.

The most striking difference between the five year needlet ILC map and the EGS3 is in the galactic plane, where the EGS3 does not recover the intermediate angular scales. The difference, however, shows no clear particular structure, as expected if it is the random realisation of a Gaussian random field. It is thus probably essentially due to the difference between our CMB reconstruction on scales larger than 3 degrees, and the CMB on larger scales that can be inferred from the CMB reconstructed by Eriksen et al. outside of their galactic mask. At higher galactic latitude, the two maps are in good agreement, with no obvious feature which could be correlated to foregrounds or to the CMB itself, with the exception of a hot spot in the large Magellanic cloud (which might be residual emission of the LMC in our map, as Eriksen et al. [76] actually mask the centre of the LMC and obtain a solution in the direction of the LMC by extrapolation from nearby pixels). Above 30 degree absolute galactic latitude, the RMS of the difference between our 3 degree map and the EGS3 map is  $5.7 \mu\text{K}$ . The two maps are in much better agreement than the EGS3 and the WMAP MEM model maps (see figure 3 in Eriksen et al.). Note however that theoretically, if there were no foregrounds in the WMAP data, the noise standard deviation  $\sigma_n$  on a 3 degree map obtained by noise-weighted averaging using all WMAP channels would be about  $3.2 \mu\text{K}$  for three year data. If instead we assumed that only the three highest frequency channels are free of foreground contamination,  $\sigma_n$  would be  $4.4 \mu\text{K}$ .

The difference between our map and the KILC5 map is more systematic with, in particular, stronger differences in the galactic plane, in spite of the fact that the two methods work on the exact same input data set. A careful visual inspection of the CMB maps themselves gives the impression that the KILC5 map is probably systematically negative towards the galactic central regions. There is, however, no secure way to be certain which map is best.

### 8.7.2 Map characterisation

#### Beam

The effective beam of the reconstructed maps are plotted in Figure 8.13, for both the full resolution five year needlet ILC map, and for the Wiener-filtered version. The map has been reconstructed for the range  $0 \leq \ell \leq 1200$ , with a smooth transition of the response, between  $\ell$  of 1024 and 1200, from the nominal W band beam value to 0. This smooth transition permits to avoid ringing effects which happen in case of a sharp cut-off in  $\ell$ . The ratio of the Wiener beam (dashed line) and ILC beam (solid line) gives a measure of the signal to noise ratio in each mode.

#### Instrumental noise

Given the ILC filter computed on the real data set, the level and properties of the instrumental noise can be straightforwardly computed by applying the *same* filter to simulated WMAP noise maps. From 100 noise realisations, we compute the average full-sky noise power spectrum (Figure 8.14), as well

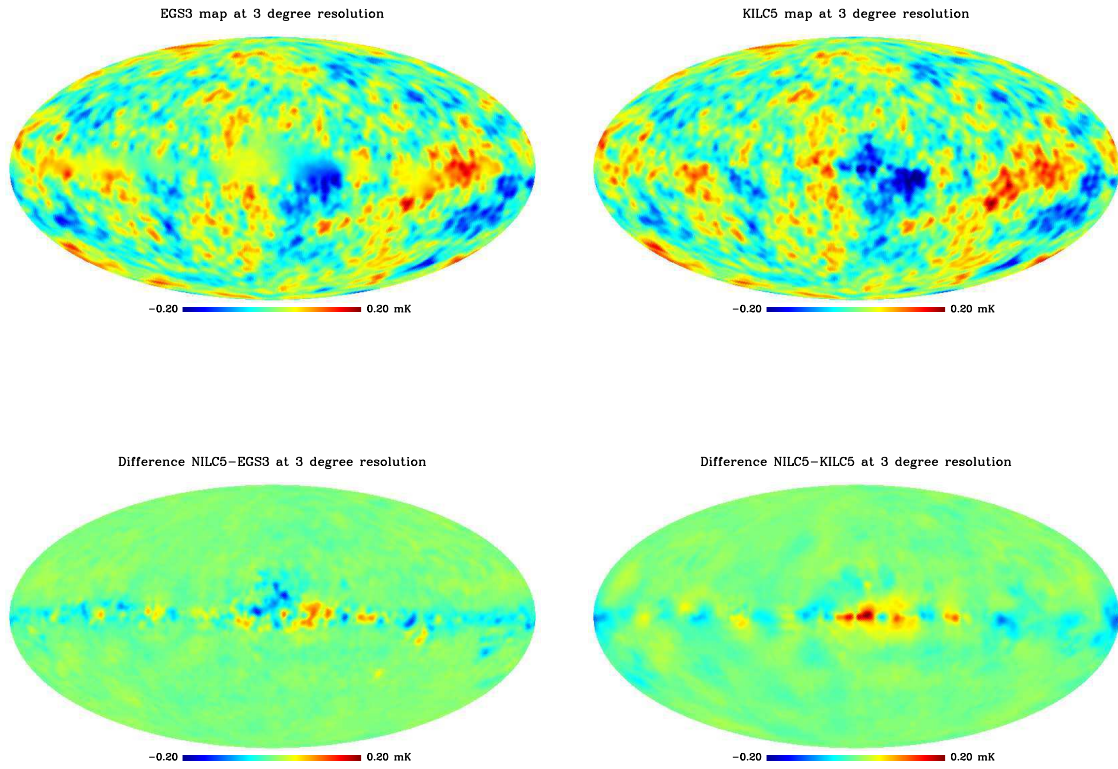


FIG. 8.12 – Top left : The EGS3 map. Top right : The KILC5 map, smoothed at 3 degree resolution. Bottom left : The difference NILC5-EGS3 ; we see is a clear difference in the galactic plane with no particular structure, compatible with a smooth Gaussian field, where the EGS3 solution poorly estimates intermediate scales ; a patch in the difference map at the location of the Large Magellanic Cloud is clearly visible also. Bottom right : The difference NILC5-KILC5 ; a clear structure aligned with the galactic plane is clearly visible, with a big difference towards the galactic center.



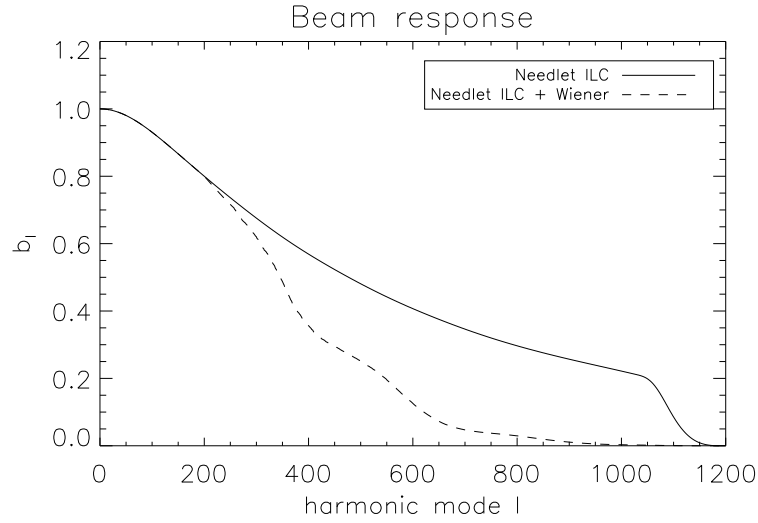


Figure 8.13 – Harmonic response of the beam of the NILC5 CMB map and of the Wiener-filtered version.

as the noise standard deviation per pixel (figure 8.15). Noise properties are not as simple as one may wish : the noise is non stationary, because of both the uneven sky coverage and of the localised processing. It is also somewhat correlated pixel-to-pixel, in particular close to the galactic plane. This is unavoidable, but fortunately our pipeline permits to make as many Monte-Carlo realisations of the instrumental noise as needed for any scientific study made using our needlet ILC map. Hundred such simulations are made available as part of our main data products.

### Foreground residuals

More problematic is the evaluation of the contamination of the CMB map by foreground residuals. It requires prior information about the foregrounds, which the ILC method avoids using. An indication of the level of systematic uncertainty is obtained from the comparison of the various solutions (figure 8.3). An other option consists in checking the contamination on simulations. Figures 8.5 and 8.6 give an idea of the expected contamination from such simulations. This, however, is good only as long as the simulations are representative enough.

The comparison of the power spectrum of the output CMB map computed at high and low galactic latitude (figure 8.10), and a visualisation of the output CMP map at high and low galactic latitude (figure 8.11) also give an idea of the amount of galactic residuals, but none of these estimates is fully satisfactory for careful CMB science. This is, however, not particular to our map. No published CMB map is provided yet with a good estimate of foreground contamination. Although this, clearly, is not fully satisfactory, we leave further investigations on this question for future work.

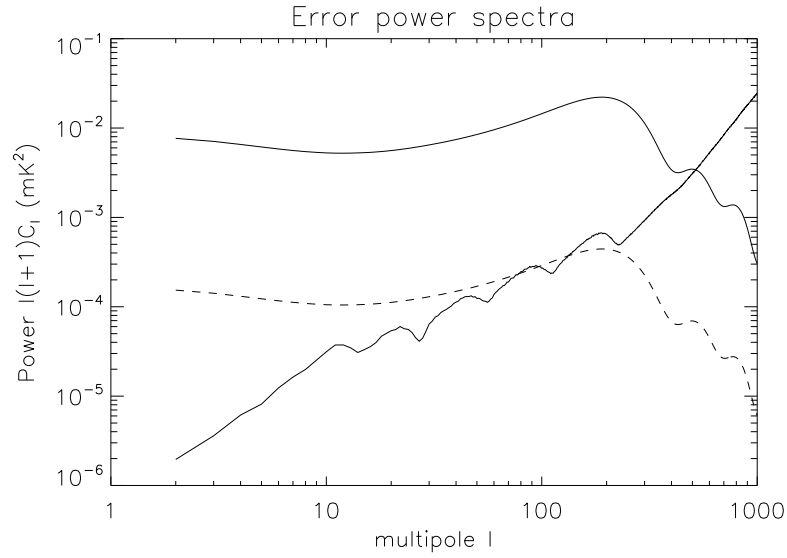


Figure 8.14 – Plot of the power spectrum of the noise (solid line). The spectrum for the WMAP best fit model is shown as a solid line also, for comparison. The dashed line is 2% of the WMAP best fit  $C_\ell$ , indicative of the level of the expected ILC bias. The bias is seen to dominate on large scales. There is, however, little margin for improvement, as few independent modes (or needlet coefficients) are available on the largest scales.

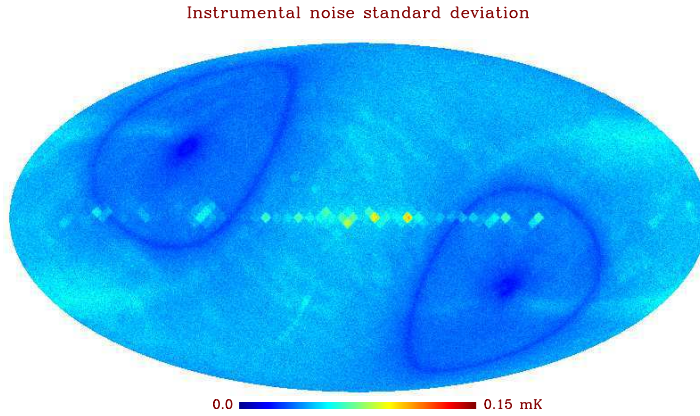


Figure 8.15 – Map of the standard deviation of the noise, per pixel at  $n_{\text{side}}=512$ , for the five year needlet ILC map at the resolution of the W channel. The adaptation of the filter to local contamination is obvious from the uniform noise level in large healpix pixels, in particular in the galactic plane. These large healpix pixels correspond to  $n_{\text{side}}=16$ , which is the size used for computing the average filter for the smallest scales. This illustrates the impact of the compromise between subtracting foregrounds and increasing the noise with the use of the lower frequency channels, which have poor sensitivity on small scales.

### 8.7.3 Final comments

#### Is our map optimal ?

In the present work, we have obtained a CMB map which has been shown to be significantly less contaminated by foregrounds and noise than the other existing maps obtained from WMAP data. A natural question is whether we can do even better.

In the following, we outline where there is margin for improvement, and explain why we have stopped the analysis at its present state.

First of all, the present analysis uses only limited external information and data sets : WMAP point source detections, and the IRIS 100 micron map. It is likely that something could be gained by using additional observations to help constraining the galactic emission.

Second, there is a trade-off between localisation of the estimation of covariance matrices, and bias in the ILC. The estimation of covariance matrices  $R_x$  over sets of  $N_p = 32 \times 32$  needlet coefficients results from a compromise which has been made based on some attempts with varying  $N_p$  on simulations, but has not been optimised in any way. In addition, the optimal solution is probably different at high galactic latitudes, where weights given to different channels probably do not vary much and thus require less localisation, and at low galactic latitudes, where the complexity of galactic emission calls for more localisation. We have tried to use  $N_p = 16 \times 16$  (more localised ILC filters, but more bias) and  $N_p = 64 \times 64$  (less localised ILC filters, but less bias). Our choice of  $N_p = 32 \times 32$  seems, on simulations, not worse than anything else (nor much better either). The bias error has been verified to remain below the reconstruction error due to the contribution of the noise for small scales, and below cosmic variance uncertainty for large scales (see figure 8.14), and remains below an acceptable level of few percent.

Similarly, the choice of the spectral window functions used on this data set has not been the object of specific optimisation. At low  $\ell$ , we follow a “dyadic” scheme, where each window reaches an  $\ell_{\max}$  about twice the previous one. Wide spectral window functions allow for more localisation in pixel space, but narrow window functions allow for more accuracy in the harmonic domain. At high  $\ell$ , because of the variation of the beams with  $\ell$ , the relative noise levels of the different channels change quite fast with  $\ell$ , which calls for more localisation in  $\ell$  space. Here again, the optimal window functions are probably not the same at high and at low galactic latitudes. In practice, we chose a small number of bands to limit the number of harmonic transforms in the pipeline and allow reasonable localisation of the analysis.

The choice we have made results from the principle of simplicity. We have tried to devise a pipeline which depends as little as possible on external data, on priors, or on fine tuning. A very simple scheme has permitted to obtain a CMB map convincingly better than other maps available. This does not preclude any attempt at more optimisation for future work if needed.

#### Why ILC and not ICA ?

It is certainly possible to tune our pipeline, changing some of its parameters. It would be possible also to use other methods than an ILC, for instance Independent Component Analysis (ICA) methods such as SMICA [69, 63], or more generally maximum likelihood methods fitting parametric models of the

foregrounds on the data sets. Such methods extract information about the foregrounds from the data directly, possibly with the help of ancillary data sets, and use this information to clean the observations in some way.

In the present case however, it is not very likely that such attempts would give much better results than what is obtained here, unless one uses a very significant number of additional data sets and safe prior information. Indeed, the WMAP data consists in five channels only, from which a component separation method based on a meaningful model of foreground emission needs to extract CMB, synchrotron, free-free, spinning dust, thermal dust (i.e. five templates), and possibly also point sources, and variations of the spectral indices of some of the components – not to mention special regions of galactic emission as cold cores and H-II regions, nor particular objects as nearby resolved galaxies or galaxy clusters. Any component separation method based on the estimation of parameters for such a rich model would almost certainly be confronted to indeterminacy issues. Methods based on a precise model are expected to be effective when the data are very redundant as compared with the number of “parameters” of the emission model, which would not be the case here. Hence, the ILC is probably one of the best approaches for doing component separation on WMAP data. It is not surprising, then, that all methods producing full sky CMB maps from WMAP, or nearly so, are implementing some variant of the ILC. Incidentally, ICA methods could benefit from the needlet framework.

## 8.8 Conclusions

In this paper, we have described a new approach for implementing CMB extraction in WMAP data, using the ILC method on a needlet frame. Tests on simulations show excellent performance of the method, thanks to localisation both in pixel space and in harmonic space. Localisation in pixel space allows the ILC weights to adapt themselves to local conditions of foreground contamination and noise (this is essentially the reason why the NILC performs better than the TILC, in particular in the vicinity of the galactic plane). Localisation in harmonic space allows to favour foreground rejection on large scales (where foregrounds dominate the total error) and instrumental noise rejection on small scales (where foregrounds are negligible but where, after beam deconvolution, the relative noise level between the various WMAP channels varies a lot as a function of scale). Needlets permit to vary the weights smoothly on large scales, and rapidly on small scales, which is not possible by cutting the sky in zones prior to any processing.

As a further improvement to previous work on WMAP data, we include a dust template in the set of analysed observations. This is motivated by the fact that on the smallest scales, observed with reasonable signal to noise ratio by the W channel only, dust emission contributes a significant fraction of the total reconstruction of the map. Using the IRIS 100 micron map as an additional observation enables the ILC to reduce the final contamination by dust –thanks to correlations of dust emission between the W channel and the 100 micron map. Special care was also taken to subtract from the data, prior to making the ILC, a number of strong point sources.

As discussed at length in the main text and in the appendix, the implementation of a filter (the ILC) based on empirical estimates of covariance matrices leads to a bias. This is not particular to our map, but is the case for any ILC map. We have estimated the level of this systematic effect, both analytically and numerically, to be at the level of a few per cent on all scales. Our simulation tool permits to make accurate estimates of the amplitude of the effect, if needed for any scientific exploitation of the NILC5 map.

The application of the method to WMAP one year and three year data (in addition to five year data) permits to compare the needlet ILC solution to previous work. Our map is seen to be at least as good as others on large scales, while being significantly less contaminated by residual foregrounds and noise than others on small scales, in particular in the vicinity of the galactic plane. The application of the method to WMAP five year data yields a CMB map which we believe to be the cleanest full sky map of the CMB to date. Contamination by noise, and the power loss due to the use of the ILC method, are characterised by means of Monte Carlo simulations. The map is available for download on the ADAMIS web site<sup>7</sup>, and can be used for a variety of science projects relying on accurate maps of the CMB.

**Acknowledgements.** The ADAMIS team at APC has been partly supported by the Astro-Map and Cosmostat ACI grants of the French ministry of research, for the development of innovative CMB data analysis methods. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science. The results in this paper have been derived using the HEALPix package [84]. The authors acknowledge the use of the Planck Sky Model<sup>8</sup> developed by the Planck working group on component separation, for making the simulations used in this work. Our pipeline is mostly implemented in octave (www.octave.org).

## 8.9 Derivation of the ILC bias

In this appendix, we compute the error made after CMB reconstruction with the ILC and some of its statistical properties. In particular, we derive the correlation of the error with the true CMB signal, which yields a non unit effective “response” of the ILC filter – and hence a *bias* in the reconstructed map and in the CMB power spectrum computed from it. This bias has to be accounted for properly for further uses of the reconstructed CMB map.

We model the data as :

$$\mathbf{x}_p = \mathbf{a}s_p + \mathbf{n}_p \quad (8.14)$$

where  $s_p$  are the coefficients of the map of interest over some domain  $\mathcal{D}$  (e.g. needlet coefficients of the CMB map for a given scale and a given patch of the sky, or pixel values in a certain region of the sky, or values of the harmonic coefficients in some band of  $\ell$ ).  $p$  indexes the coefficient (i.e. pixel, or harmonic mode, or needlet coefficient).  $\mathbf{x}_p$  are the observations of coefficient  $p$  for the set of available observed maps, and  $\mathbf{n}_p$  the corresponding “noise” terms (including foreground contaminants).

The ILC is best applied over domains of  $p$  where all coefficient have (near) uniform expected signal and noise properties, so that the ILC weights are optimal simultaneously for all  $p$ . In particular, the RMS values of all maps do not depend (much) on  $p$  in a given domain. Hence, we concentrate on one given domain of  $p$  for which we assume that the sequences  $s_p$  and  $\mathbf{n}_p$  are independent, Gaussian random variables with distribution  $\mathcal{N}(0, \sigma_s^2)$  and  $\mathcal{N}(0, \mathbf{R}_n)$ , with  $\sigma_s^2$  the variance of  $s_p$  (the CMB) and  $\mathbf{R}_n$  the covariance matrix of the noise (including foregrounds).

<sup>7</sup>[http://www.apc.univ-paris7.fr/APC\\_CS/Recherche/Adamis/cmb\\_wmap-en.php](http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/cmb_wmap-en.php)

<sup>8</sup>[http://www.apc.univ-paris7.fr/APC\\_CS/Recherche/Adamis/PSM/psky-en.php](http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/PSM/psky-en.php)

The ILC estimate of  $s_p$  in domain  $\mathcal{D}$  is given, for all  $p$ , by

$$\hat{s}_p = \frac{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1}}{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{a}} \mathbf{x}_p \quad (8.15)$$

where  $\widehat{\mathbf{R}}_x$  is an estimate of the covariance matrix of the observations, obtained as :

$$\begin{aligned} \widehat{\mathbf{R}}_x &= \frac{1}{N_p} \sum_p \mathbf{x}_p \mathbf{x}_p^t \\ &= \frac{1}{N_p} \sum_p (\mathbf{a} s_p + \mathbf{n}_p)(\mathbf{a} s_p + \mathbf{n}_p)^t \end{aligned} \quad (8.16)$$

with  $N_p$  the number of coefficients in domain  $\mathcal{D}$ . In the limit of large  $N_p$ ,  $\widehat{\mathbf{R}}_x$  approaches its expectation (ensemble average) value  $E(\widehat{\mathbf{R}}_x) = \mathbf{R}_x$ . For finite  $N_p$ , we have instead

$$\widehat{\mathbf{R}}_x = \mathbf{R}_x + \Delta_x \quad (8.17)$$

where  $\Delta_x$  is a correction term corresponding to the departure of the empirical correlation from its ensemble average due to the finite sample size  $N_p$ . From now on, we assume that  $N_p$  is large enough that this correction is small : we investigate effects at first order in  $1/N_p$ .

### 8.9.1 First order expansion of the reconstruction error

We are interested in the reconstruction error :

$$d_p = \hat{s}_p - s_p \quad (8.18)$$

The first and second moments of  $d_p$ , i.e. the mean value  $E(d_p)$  of the reconstruction error, as well as its variance  $E(d_p^2)$ , are of particular interest for the interpretation of the reconstructed map. In particular, we have

$$E(\hat{s}_p^2) = E(s_p^2) + E(d_p^2) + 2E(s_p d_p) \quad (8.19)$$

In our case,  $E(\hat{s}_p^2)$  can be used to estimate  $E(s_p^2)$ . In the case where domain  $\mathcal{D}$  is an harmonic domain,  $p$  indexes harmonic modes  $(\ell, m)$ , and  $E(s_p^2)$  is a term of the CMB power spectrum. In our needlet approach,  $E(s_p^2)$  is also directly connected to the CMB power spectrum. For this reason, it is important to characterise in the best way we can the “noise bias”  $E(d_p^2)$  and the covariance of the error with the CMB,  $E(s_p d_p)$ .

The ILC being constructed so that the response to the signal of interest is unity, only the filtered noise term contributes to the error  $d_p$ , which can then be written as :

$$d_p = \frac{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1}}{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{a}} \mathbf{n}_p \quad (8.20)$$

$$= \frac{\mathbf{a}^t [\mathbf{R}_x + \Delta_x]^{-1}}{\mathbf{a}^t [\mathbf{R}_x + \Delta_x]^{-1} \mathbf{a}} \mathbf{n}_p \quad (8.21)$$

where  $\Delta_x$  is a small perturbation to  $\mathbf{R}_x$ . We use the first order expansion :

$$[\mathbf{R} + \Delta_x]^{-1} \simeq \mathbf{R}^{-1} - \mathbf{R}^{-1} \Delta_x \mathbf{R}^{-1} \quad (8.22)$$

which yields

$$d_p = \frac{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{n}_p}{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{a}}$$

Writing :

$$\begin{aligned} \frac{1}{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{a}} &= \frac{1}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \frac{1}{(1 - \epsilon)} \\ &\simeq \frac{1}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} (1 + \epsilon) \end{aligned}$$

where  $\epsilon$  is

$$\epsilon = \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{a}}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}}$$

we get

$$d_p \simeq \frac{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} (1 + \epsilon)$$

Keeping only first order terms in  $(\Delta_x)$  yields :

$$d_p = \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} - \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} + \frac{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p][\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{a}]}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2} \quad (8.23)$$

The first term on the right hand side, proportional to  $\mathbf{n}$ , is the “ideal” ILC error, i.e. the error we would get if we knew perfectly the “true” covariance matrix  $\mathbf{R}_x$  of the observations. The second and third terms, proportional to  $\Delta_x$ , are corrections due to the fact that this covariance matrix is actually obtained empirically from the observation themselves.

From equations (8.16) and (8.17), we can write  $\Delta_x$  in the form :

$$\Delta_x = \delta_s \mathbf{a} \mathbf{a}^t + \Delta_n + \widehat{\mathbf{C}}$$

where

$$\begin{aligned} \delta_s &= \widehat{\sigma}_s^2 - \sigma_s^2 \\ \Delta_n &= \widehat{\mathbf{R}}_n - \mathbf{R}_n \\ \widehat{\mathbf{C}} &= \frac{1}{N_p} \sum_q (\mathbf{n}_q s_q \mathbf{a}^t + \mathbf{a} s_q \mathbf{n}_q^t) \end{aligned}$$

These three quantities correspond respectively, in pixel (or mode, or needlet coefficient)  $p$ , to the uncertainty in CMB variance estimates due to “cosmic” (or sample) variance, to the error in the estimation of the “noise” covariance matrix alone (if maps of noise+foregrounds alone were available), and to a cross term, originating from the empirical covariance between CMB and noise due the finite sample size  $N_p$ .

The two last terms (small correction terms) in equation 8.23, being proportional to  $\Delta_x$ , can be decomposed each into the sum of three terms, proportional to  $(\delta_s \mathbf{a} \mathbf{a}^t)$ ,  $\Delta_n$ , and  $\widehat{\mathbf{C}}$  respectively. The signal and noise realisations enter the  $(\delta_s \mathbf{a} \mathbf{a}^t)$  term as products of terms of the form  $(\mathbf{a} \mathbf{a}^t s_q s_q \mathbf{n}_p)$  only, the  $\Delta_n$  term as products of terms of the form  $(\mathbf{n}_q \mathbf{n}_q^t \mathbf{n}_p)$ . On the contrary, signal and noise realisations enter

the  $\widehat{\mathbf{C}}$  term as the product of terms in the form  $(\mathbf{a}s_q\mathbf{n}_q^t\mathbf{n}_p)$ , i.e. products of  $s$  and second power of  $\mathbf{n}$ . Index  $q$  runs over domain  $\mathcal{D}$ .

Assuming that  $s$  and  $\mathbf{n}$  are centred variables, the mean value of the error is immediately seen to vanish :

$$\mathbb{E}(d_p) = 0$$

The main contribution to the variance comes from the first term on the right hand side of eq. 8.23. The second and third terms are small corrections to this variance estimate, so that to first order, we get :

$$\mathbb{E}(d_p^2) \simeq \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{R}_n \mathbf{R}_x^{-1} \mathbf{a}}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2}$$

Recalling that  $\mathbf{R}_x = \mathbf{R}_n + \sigma^2 \mathbf{a} \mathbf{a}^t$ , where  $\sigma^2$  is the variance of the CMB, and making use of the inversion formula :

$$[\mathbf{R}_n + \sigma^2 \mathbf{a} \mathbf{a}^t]^{-1} = \mathbf{R}_n^{-1} - \sigma^2 \frac{\mathbf{R}_n^{-1} \mathbf{a} \mathbf{a}^t \mathbf{R}_n^{-1}}{1 + \sigma^2 \mathbf{a}^t \mathbf{R}_n^{-1} \mathbf{a}}$$

we finally obtain :

$$\mathbb{E}(d_p^2) \simeq \frac{1}{[\mathbf{a}^t \mathbf{R}_n^{-1} \mathbf{a}]}$$

The most interesting terms are those connecting the error to the signal of interest<sup>9</sup>,  $\mathbb{E}(s_p d_p)$ , which is necessary to compute the power spectrum of the output map according to 8.19.

As mentioned previously, under the assumption that the signal of interest is not correlated to the noise and the foregrounds, the first term (main term) of the r.h.s. of equation 8.23 does not give rise to multiplicative errors (or correlation of  $d_p$  with  $s_p$ ). Similarly, the corrective term proportional to  $\delta_s \mathbf{a} \mathbf{a}^t + \Delta_n$ , multiplied by  $s$ , gives terms which contain an odd power of  $s$  and an odd power of  $\mathbf{n}$ , and does not give rise to correlations. This assumption is correct, to excellent accuracy, when the signal of interest is CMB anisotropies<sup>10</sup>. We are left with :

$$\begin{aligned} \mathbb{E}(s_p d_p) = & \mathbb{E} \left( \sum_q \frac{s_p s_q}{N_p} \frac{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p] [\mathbf{a}^t \mathbf{R}_x^{-1} (\mathbf{n}_q \mathbf{a}^t + \mathbf{a} \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{a}]}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2} \right) \\ & - \mathbb{E} \left( \sum_q \frac{s_p s_q}{N_p} \frac{\mathbf{a}^t \mathbf{R}_x^{-1} (\mathbf{n}_q \mathbf{a}^t + \mathbf{a} \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \right) \end{aligned}$$

Multiplying the numerator and denominator of the second term by  $\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}$  and expanding numerators, two terms cancel and two remain. If we assume in addition that signal and/or noise coefficients are

<sup>9</sup>We warn the reader that some authors fail to make a clear distinction between the statistical (ensemble average) correlation, which is a deterministic quantity, and the “empirical correlations” computed, assuming some kind of ergodicity, as averages over finite sets of samples as in equation 8.16.

<sup>10</sup>Certainly the CMB is not correlated to galactic components. Small correlations with large scale structure, and hence with SZ effect and emission from outer galaxies, may exist because of the integrated Sachs-Wolfe effect. We neglect this effect in the present discussion.



independent, i.e.  $E(s_p s_q) = \sigma_s^2 \delta_{qp}$  and/or  $E(\mathbf{n}_p \mathbf{n}_q^t) = \mathbf{R}_n \delta_{qp}$ , only the  $pp$  term is non vanishing, and we get

$$E(s_p d_p) = \frac{\sigma_s^2}{N_p} \left( \frac{E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)^2) - (\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}) E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_p)}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} \right) \quad (8.24)$$

We compute

$$\begin{aligned} E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)^2) &= \mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{R}_n \mathbf{R}_x^{-1} \mathbf{a} \\ &= \mathbf{a}^t \mathbf{R}_x^{-1} [\mathbf{R}_x - \sigma_s^2 \mathbf{a} \mathbf{a}^t] \mathbf{R}_x^{-1} \mathbf{a} \\ &= [\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}] [1 - \sigma_s^2 \mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}] \end{aligned} \quad (8.25)$$

and

$$\begin{aligned} E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_p) &= \text{Tr}(\mathbf{R}_x^{-1} \mathbf{R}_n) \\ &= \text{Tr}(\mathbf{R}_x^{-1} [\mathbf{R}_x - \sigma_s^2 \mathbf{a} \mathbf{a}^t]) \\ &= \text{Tr}(\text{Id}) - \sigma_s^2 \text{Tr}(\mathbf{R}_x^{-1} \mathbf{a} \mathbf{a}^t) \\ &= m - \sigma_s^2 (\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}) \end{aligned} \quad (8.26)$$

where  $m$  is the number of channels used for the ILC (here, 5 WMAP channels + 1 IRIS map, for a total of 6). Substituting the results of equations 8.25 and 8.26 into equation 8.24, we get the simple final result :

$$E(s_p d_p) = \frac{\sigma_s^2 (1 - m)}{N_p} \quad (8.27)$$

The error in the reconstructed CMB map comprises a term proportional (on average) to the CMB. In our application,  $m = 6$  and  $N_p = 1024$ , so that if indeed all needlet coefficients were independent, the amplitude of the effect should be  $E(s_p d_p) \simeq 5 \times 10^{-3} \sigma_s^2$ , i.e. a bias of about half a percent in the CMB reconstruction.

### 8.9.2 A geometric interpretation

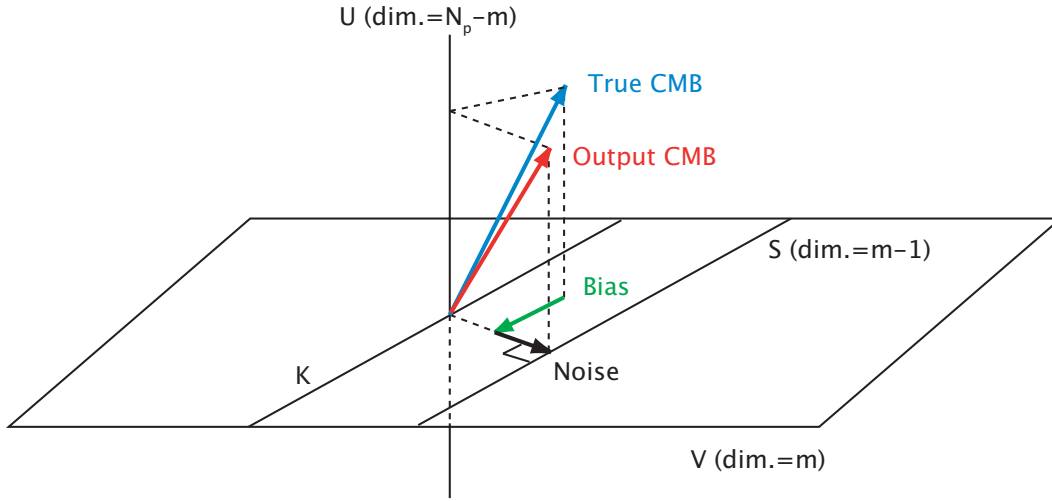
Although allowing the statistical derivation of the (anti-) correlation of the reconstruction error with the original CMB, the above derivation is not very illuminating about the mechanism giving rise to this CMB power loss. A geometrical reasoning gives better insight on what is actually going on.

For a given data set, the ILC works on one single realisation of all random fields. For an independent implementation of the ILC on  $N_p$  pixels (or modes, or needlet coefficients) of the observations, each data set is represented by a vector in a  $N_p$ -dimensional vector space  $W$ . The CMB  $\mathbf{s}$ , the observation  $\mathbf{x}_i$  for each channel  $i$ , and each of the noise realisations  $\mathbf{n}_i$  (including foregrounds) are elements of  $W$ .

The collection of vectors  $\mathbf{n}_i$  defines an  $m$ -dimensional subspace  $V$  of  $W$ . This is true irrespective of the nature of the foregrounds : indeed, although in principle vectors  $\mathbf{n}_i$  could be linearly dependent, this happens in practice with vanishing probability (in particular if the observations are noisy).

Vector space  $W$  can thus be decomposed in two orthogonal subspaces,  $U$  and  $V$ , where  $V$  is the  $m$ -dimensional sub-space spanned by all vectors  $\mathbf{n}_i$ , and  $U = V^\perp$  is a  $(N_p - m)$  dimensional vector space. The CMB itself can be decomposed into two components, one in  $U$ , and one in  $V$  :

$$\mathbf{s} = \mathbf{s}_U + \mathbf{s}_V$$



F . 8.16 – Geometric illustration of the ILC bias. The true CMB (blue) and the output CMB (red) are elements of  $W = U \oplus V$ . The difference between the two can be decomposed into the sum of two elements of  $V$  : a bias (green), which is an element of  $K$ , and a noise contribution (black), which is an element of  $H$ , the orthogonal of  $K$  in  $V$ .

where  $s_U$  is the orthogonal projection of  $s$  onto  $U$ , and  $s_V$  its orthogonal projection onto  $V$ .

What happens when the ILC is made is the following : we look for weights  $w_i$  for all channels, that minimise the variance of the reconstructed map, i.e. minimise the norm of vector  $\widehat{s} = \sum_i w_i \mathbf{x}_i$ , under the constraint that  $\sum_i w_i = 1$ . We have :

$$\begin{aligned} \widehat{s} &= s_U + \left( s_V + \sum_i w_i \mathbf{n}_i \right) \\ &= \widehat{s}_U + \widehat{s}_V \end{aligned}$$

where the first term is a vector of  $U$  and the second term a vector of  $V$ , and the second line of the equation defines  $\widehat{s}_U$  and  $\widehat{s}_V$ . Since these two subspaces of  $W$  are orthogonal, the norm of  $\widehat{s}$  is the sum of the norms of the two vectors  $\widehat{s}_U$  and  $\widehat{s}_V$ . The norm of  $\widehat{s}$  thus depends on  $w_i$  only through the norm of the projection of  $\widehat{s}$  on subspace  $V$ .

The noise contribution to  $\widehat{s}$  appears as a linear combination of vectors  $\mathbf{n}_i$ . For varying values of  $w_i$  such that  $\sum_i w_i = 1$ , this linear combination spans an *affine* subspace  $S$  of  $V$ .  $S$  is of dimension  $m - 1$  (an hyperplane). Defining  $K$  as the vector subspace of  $V$  spanned by linear combinations  $\sum_i w_i \mathbf{x}_i$  such that  $\sum_i w_i = 0$ , we obtain  $S$  as :

$$S = \mathbf{p} + K$$

where  $\mathbf{p}$  is any element of  $S$ .

We note that the vector subspace  $K$  depends only on noise realisations, and not on  $s$  nor on the final ILC weights (the latter only defining a single point on  $S$  – and on  $K$  by orthogonal projection). Hence, the direction of the one-dimensional vector subspace  $H$  of  $V$  orthogonal to  $K$  is also independent of  $s$  and of the final ILC weights.

For any element  $\widehat{s}_V = s_V + \sum_i w_i \mathbf{n}_i$  of affine space  $S$  the norm of  $\widehat{s}_V$  is the sum of the norms of its projections onto  $K$  and  $H$ . Allowing weights  $w_i$  to vary, vector  $\widehat{s}_V$  spans  $S$ , and hence only the norm

of the projection of  $\widehat{s}_V$  onto  $K$  varies (and not its projection on  $H$ ). The minimum is reached when the projection of  $\widehat{s}_V$  onto  $K$  vanishes. When this happens, the ILC has cancelled completely the linear combination of projections of  $s$  and  $\mathbf{n}_i$  onto the  $m - 1$ -dimensional space  $K$ , and left untouched the projections of  $s$  and  $\mathbf{n}_i$  onto the  $N_p - m + 1$  dimensional vector space  $U \oplus H$ .

Assuming the CMB to be Gaussian and statistically isotropic, its coefficients in any orthogonal basis are Gaussian distributed random variables with variance  $\sigma^2/N_p$  (since the sum must have total variance  $\sigma^2$ ). It follows straightforwardly that the correlation of the recovered CMB map with the input CMB map is  $(N_p - m + 1)/N_p$ , and that the “bias” is due to the loss of  $m - 1$  modes of the original CMB, which have been unlucky enough to “live” in the  $(m - 1)$  dimensional space  $K$ .

### 8.9.3 Comment on coefficient independence

The above derivation in Section 8.9.1 assumes the independence of coefficients  $\mathbf{n}_p$  and/or of coefficients  $s_p$ , i.e.  $E(\mathbf{n}_p \mathbf{n}_q^t) = \mathbf{R}_n \delta_{qp}$  and/or  $E(s_p s_q) = \sigma_s^2 \delta_{qp}$ . When this assumption does not hold, we have :

$$E(s_p d_p) = \frac{1}{N_p} \sum_q E \left( s_p s_q \frac{(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q)}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} \right) - \frac{1}{N_p} \sum_q E(s_p s_q) E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q)$$

Assuming that the noise and the CMB are independent, we have :

$$E(s_p d_p) = \frac{1}{N_p} \sum_q E(s_p s_q) \frac{E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q))}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} - \frac{1}{N_p} \sum_q E(s_p s_q) E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q) \quad (8.28)$$

where

$$E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q)) = \mathbf{a}^t \mathbf{R}_x^{-1} E(\mathbf{n}_p \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{a} \quad (8.29)$$

and

$$E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q) = \text{Tr}(\mathbf{R}_x^{-1} E(\mathbf{n}_p \mathbf{n}_q^t)) \quad (8.30)$$

When  $p$  and  $q$  index needlet coefficients as in the present work, we have :

$$E(s_p s_q) = \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_{\ell}^2 C_{\ell} P_{\ell}(\cos \theta_{qp})$$

where  $s_p$  and  $s_q$  are needlet coefficients of the CMB map, evaluated at two different points  $p$  and  $q$ ,  $C_{\ell}$  is the angular power spectrum of the CMB,  $\theta_{qp}$  is the angle between  $q$  and  $p$ , and  $N_{\text{tot}}$  is the total number of pixels of the needlet coefficient map. For noise maps (including foregrounds), which are not stationary Gaussian random fields on the sphere, the analogous formula is just an approximation, which can be written as :

$$E(\mathbf{n}_p \mathbf{n}_q^t) \simeq \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_{\ell}^2 \mathbf{R}_n(\ell) P_{\ell}(\cos \theta_{qp})$$

where  $\mathbf{n}_p$  and  $\mathbf{n}_q$  are needlet coefficients of all noise maps, evaluated at two different points  $p$  and  $q$ ,  $\mathbf{R}_n(\ell)$  is the covariance of the noise needlet coefficients (an  $m \times m$  matrix for each  $\ell$ ), and  $\theta_{qp}$  and  $N_{\text{tot}}$  are defined as above.

Assuming that neither the noise level, nor the CMB power, do vary much over the spectral window  $h_\ell$ ,  $R_n(\ell)$  and  $C_\ell$  are approximately independent of  $\ell$ , and can be taken out of the integral. We get :

$$E(s_p s_q) = C_\ell k(\theta_{qp})$$

and

$$E(\mathbf{n}_p \mathbf{n}_q^t) = R_n k(\theta_{qp})$$

with

$$k(\theta_{qp}) = \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_\ell^2 P_\ell(\cos \theta_{qp})$$

Hence, plugging this result, together with equations 8.29 and 8.30, into equation 8.28 we get :

$$E(s_p d_p) = \frac{(1 - m)}{N_p} C_\ell \sum_q k^2(\theta_{qp})$$

Finally, noting that  $\sigma_s^2 = C_\ell k(0)$ , we get

$$E(s_p d_p) = \frac{(1 - m)}{N_p} \frac{\sigma_s^2}{k(0)} \sum_q k^2(\theta_{qp}) \quad (8.31)$$

Equation 8.31 is the exact same as equation 8.27, except for a coefficient, which measures the correlation between signal and noise coefficients  $p$  and coefficients  $q$  in domain  $\mathcal{D}$ . In particular, the result is, again, independent of  $R_n$ .

Hence, we define an *effective number of modes*,  $N_p^{\text{eff}} = N_p/\alpha$ , where

$$\alpha = \frac{\sum_q k^2(\theta_{qp})}{k(0)}$$

and we get

$$E(d_p s_p) = \frac{\sigma_p^2 (1 - m)}{N_p^{\text{eff}}}$$

An approximation (and upper bound) for  $\alpha$  is easily obtained in the special case where  $h_\ell$  is a square spectral window, and when domains  $\mathcal{D}$  over which the ILC is implemented are small regions of the sky, so that  $h_\ell^2 = 1$ , and  $P_\ell(\cos \theta_{qp}) \simeq 1$ . In this case, we have :

$$k(\theta_{qp}) \simeq \frac{1}{N_{\text{tot}}} \left( (\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right)$$

and we have

$$N_p^{\text{eff}} \simeq \frac{N_p}{N_{\text{tot}}} \left( (\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right) \quad (8.32)$$

We note that  $\left( (\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right)$  simply is the number of harmonic coefficients selected by the spectral window  $h_\ell$ , and that  $N_p/N_{\text{tot}} = f_{\text{sky}}$  is a coefficient which takes into account the effect of partial sky coverage for the (local) calculation of the statistics of the data set, well in line with  $N_p^{\text{eff}}$  being understood as an effective number of modes.

# Bibliographie

*N.B. Les travaux qui suivent sont regroupés selon le domaine d'intérêt principal du journal qui les a publiés. Néanmoins, de nombreux articles seraient à leur place dans plusieurs de ces catégories.*

## Analyse et Mathématiques générales

- [1] J.-P. Antoine, S. Dahlke, and M. Holschneider (Eds). Special Issue : Analysis on the Sphere I. *Journal of Fourier Analysis and Applications*, 13(4) :369–510, 2007.
- [2] J.-P. Antoine, S. Dahlke, and M. Holschneider (Eds). Special Issue : Analysis on the Sphere II. *Journal of Fourier Analysis and Applications*, 13(6) :643–748, 2007.
- [3] J.-P. Antoine, L. Demanet, L. Jacques, and P. Vandergheynst. Wavelets on the sphere : implementation and approximations. *Applied and Computational Harmonic Analysis*, 13(3) :177–200, 2002.
- [4] J.-P. Antoine and P. Vandergheynst. Wavelets on the n-sphere and related manifolds. *Journal of Mathematical Physics*, 39(8) :3987–4008, 1998.
- [5] J.-P. Antoine and P. Vandergheynst. Wavelets on the 2-sphere : a group-theoretical approach. *Applied and Computational Harmonic Analysis*, 7(3) :262–291, 1999.
- [6] I. Bogdanova, P. Vandergheynst, J.-P. Antoine, L. Jacques, and M. Morvidone. Stereographic wavelet frames on the sphere. *Applied and Computational Harmonic Analysis*, 19(2) :223–252, 2005.
- [7] S. Dahlke, W. Dahmen, E. Schmitt, and I. Weinreich. Multiresolution analysis and wavelets on  $S^2$  and  $S^3$ . *Numerical Functional Analysis and Optimization*, 16(1-2) :19–41, 1995.
- [8] S. Dahlke and P. Maass. Continuous wavelet transforms with applications to analyzing functions on spheres. *Journal of Fourier Analysis and Applications*, 2(4) :379–396, 1996.
- [9] I. Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1992.
- [10] S. de Bièvre. Cohérent states over symplectic homogeneous spaces. *Journal of Mathematical Physics*, 30(3) :1401–1407, 1989.

- [11] V. Drinfel'd. Finitely additive measures on  $S^2$  and  $S^3$ , invariant with respect to rotations. *Functional Analysis and its Applications*, 18 :245–246, 1984.
- [12] M. Elad, J.-L. Starck, D. Donoho, and P. Querre. Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA). *Applied and Computational Harmonic Analysis*, 19 :340–358, 2005.
- [13] N. Fernández. Optimally space-localized band-limited wavelets on  $\mathbb{S}^{q-1}$ . *Journal of Computational and Applied Mathematics*, 199(1) :68–79, 2007.
- [14] W. Freeden, T. Gervens, and M. Schreiner. *Constructive Approximation on the Sphere*. Clarendon Press, 1998.
- [15] W. Freeden and V. Michel. *Multiscale potential theory. With applications to geoscience*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2004.
- [16] W. Freeden and F. Schneider. Regularization wavelets and multiresolution. *Inverse Problems*, 14(2) :225–243, 1998.
- [17] W. Freeden and M. Schreiner. Non-orthogonal expansions on the sphere. *Mathematical Methods in the Applied Sciences*, 18 :83–120, 1995.
- [18] W. Freeden and M. Schreiner. Orthogonal and nonorthogonal multiresolution analysis, scale discrete and exact fully discrete wavelet transform on the sphere. *Constructive Approximation*, 14(4) :493–515, 1998.
- [19] W. Freeden and U. Windheuser. Spherical wavelet transform and its discretization. *Advances in Computational Mathematics*, 5(1) :51–94, 1996.
- [20] W. Freeden and U. Windheuser. Combined spherical harmonic and wavelet expansion – A future concept in Earth's gravitational determination. *Applied and Computational Harmonic Analysis*, 4 :1–37, 1997.
- [21] D. Geller and A. Mayeli. Nearly tight frames and space-frequency analysis on compact manifolds. *Mathematische Zeitschrift*, 2008. (In press).
- [22] S. Goh and T. Goodman. Uncertainty principles and asymptotic behavior. *Applied and Computational Harmonic Analysis*, 16(1) :19–43, 2004.
- [23] J. Göttelmann. Locally supported wavelets on manifolds with applications to the 2d sphere. *Applied and Computational Harmonic Analysis*, 7 :1–33, 1999.
- [24] F. Grünbaum, L. Longhi, and M. Perlstadt. Differential operators commuting with finite convolution integral operators : some nonabelian examples. *SIAM Journal on Applied Mathematics*, 42(5) :941–955, 1982.
- [25] F. Guilloux, G. Faÿ, and J.-F. Cardoso. Practical wavelet design on the sphere. *Applied and Computational Harmonic Analysis*, 2008. (In press).
- [26] M. Holschneider. Continuous wavelet transforms on the sphere. *Journal of Mathematical Physics*, 37(8) :4156–4165, 1996.

- [27] T. MacRobert. *Spherical harmonics. An elementary treatise on harmonic functions with applications*. Third edition. International Series of Monographs in Pure and Applied Mathematics, Vol. 98. Pergamon Press, Oxford, 1967.
- [28] J. McEwen, P. Vielva, Y. Wiaux, R. Barreiro, I. Cayón, M. Hobson, A. Lasenby, E. Martínez-González, and J. Sanz. Cosmological applications of a wavelet analysis on the sphere. *Journal of Fourier Analysis and Applications*, 13(4) :495–510, 2007.
- [29] H. Mhaskar, F. Narcowich, and J. Ward. Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature. *Mathematics of Computation*, 70(235) :1113–1130, 2001.
- [30] I. Moore and M. Cada. Prolate spheroidal wave functions, an introduction to the slepian series and its properties. *Applied and Computational Harmonic Analysis*, 16 :208–230, 2004.
- [31] C. Müller. *Spherical harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer-Verlag, 1966.
- [32] F. Narcowich, P. Petrushev, and J. Ward. Decomposition of Besov and Triebel-Lizorkin spaces on the sphere. *Journal of Functional Analysis*, 238 :530–564, 2006.
- [33] F. Narcowich, P. Petrushev, and J. Ward. Localized tight frames on spheres. *SIAM Journal on Mathematical Analysis*, 38(2) :574–594, 2006.
- [34] F. Narcowich and J. Ward. Nonstationary wavelets on the  $m$ -sphere for scattered data. *Applied and Computational Harmonic Analysis*, 3 :324–336, 1996.
- [35] F. Narcowich and J. Ward (Eds). Special Issue : Approximation, computation and representation of functions on the sphere. *Advances in Computational Mathematics*, 21(1), 2004.
- [36] F. Narcowich and J. Ward (Eds). Special Issue : Approximation, computation and representation of functions on the sphere. *Advances in Computational Mathematics*, 21(2), 2004.
- [37] D. Potts, G. Steidl, and M. Tasche. Kernels of spherical harmonics and spherical frames. In F. Fontanella, K. Jetter, and P. J. Laurent, editors, *Advances Topics in Multivariate Approximation*, pages 1–154. World Scientific Publishing, 1996.
- [38] D. Potts and M. Tasche. Interpolatory wavelets on the sphere. In C. Chui and L. Schumaker, editors, *Approximation Theory VIII, Vol. 2 : Wavelets*, pages 335–342. World Scientific, 1995.
- [39] P. Schröder and W. Sweldens. Spherical wavelets : Efficiently representing functions on the sphere. In *Computer Graphics Proceedings (SIGGRAPH 95)*, pages 161–172, 1995.
- [40] R. Seeley. Spherical harmonics. *The American Mathematical Monthly*, 73(4, part II) :115–121, 1966.
- [41] F. Simons, F. Dahlen, and M. Wieczorek. Spatiospectral concentration on a sphere. *SIAM Review*, 48(504), 2006.
- [42] D. Slepian. Some comments on Fourier analysis, uncertainty and modeling. *SIAM Review*, 25(3) :379–393, 1983.
- [43] E. Stein and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton Mathematical Series, No. 32. Princeton University Press, 1971.

- [44] J. Talman. *Special Functions : A group theoretical approach*. Benjamin, 1968.
- [45] B. Torresani. *Analyse continue par ondelettes*. 1re édition. Savoirs Actuels. InterÉditions / CNRS Éditions, 1995.
- [46] B. Torresani. Position-frequency analysis for signals defined on spheres. *Signal Processing*, 43(3) :341–346, 1995.
- [47] G. Walter and X. Shen. Sampling with prolate spheroidal wave functions. *Sampling Theory in Signal and Image Processing*, 2(1) :25–52, 2003.
- [48] G. Walter and X. Shen. Wavelets based on prolate spheroidal wave functions. *Journal of Fourier Analysis and Applications*, 10(1) :1–26, 2004.
- [49] G. Walter and T. Soleski. A new friendly method of computing prolate spheroidal wave functions and wavelets. *Applied and Computational Harmonic Analysis*, 19 :432–443, 2005.
- [50] Y. Wiaux, L. Jacques, and P. Vandergheynst. Correspondence principle between spherical and euclidean wavelets. *The Astrophysical Journal*, 632 :15, 2005.
- [51] M. Wiecezorek and F. Simons. Minimum-variance multitaper spectral estimation on the sphere. *Journal of Fourier Analysis and Applications*, 13(6) :665–692, 2007.

## Cosmologie et Astrophysique

- [52] R. Aurich, S. Lustig, and F. Steiner. The circles-in-the-sky signature for three spherical universes. *Monthly Notices of the Royal Astronomical Society*, 369 :240–248, 2006.
- [53] C. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. Meyer, L. Page, D. Spergel, G. Tucker, E. Wollack, E. Wright, C. Barnes, M. Greason, R. Hill, E. Komatsu, M. Nolte, N. Odegard, H. Peiris, L. Verde, and J. Weiland. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Preliminary Maps and Basic Results. *The Astrophysical Journal*, 148 :1–27, 2003.
- [54] C. Bennett, R. Hill, G. Hinshaw, M. Nolte, N. Odegard, L. Page, D. Spergel, J. Weiland, E. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. Meyer, G. Tucker, and E. Wollack. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Foreground Emission. *The Astrophysical Journal Supplement Series*, 148 :97–117, 2003.
- [55] A. Benoît, P. Ade, A. Amblard, R. Ansari, E. Aubourg, J. Bartlett, J.-P. Bernard, R. Bhatia, A. Blanchard, J. Bock, A. Boscaleri, F. Bouchet, A. Bourrachot, P. Camus, F. Couchot, P. de Bernardis, J. Delabrouille, F.-X. Désert, O. Doré, M. Douspis, L. Dumoulin, X. Dupac, P. Filliatre, K. Ganga, F. Gannaway, B. Gautier, M. Giard, Y. Giraud-Héraud, R. Gispert, L. Guiglielmi, J.-C. Hamilton, S. Hanany, S. Henrot-Versillé, V. Hristov, J. Kaplan, G. Lagache, J.-M. Lamarre, A. Lange, K. Madet, B. Maffei, D. Marrone, S. Masi, J. Murphy, F. Naraghi, F. Nati, G. Perrin, M. Piat, J.-L. Puget, D. Santos, R. Sudiwala, J.-C. Vanel, D. Vibert, E. Wakui, and D. Yvon. Archeops : a high resolution, large sky coverage balloon experiment for mapping cosmic microwave background anisotropies. *Astroparticle Physics*, 17 :101–124, 2002.



- [56] A. Bonaldi, L. Bedini, E. Salerno, C. Baccigalupi, and G. de Zotti. Estimating the spectral indices of correlated astrophysical foregrounds by a second-order statistical approach. *Monthly Notices of the Royal Astronomical Society*, 373 :271–279, 2006.
- [57] A. Bonaldi, S. Ricciardi, S. Leach, F. Stivoli, C. Baccigalupi, and G. De Zotti. WMAP 3yr data with the CCA : anomalous emission and impact of component separation on the CMB power spectrum. *preprint*, 2007.
- [58] J. Bond, A. Jaffe, and L. Knox. Estimating the power spectrum of the cosmic microwave background. *Physical Review D*, 57(4) :2117–2137, 1998.
- [59] P. Cabella, F. K. Hansen, M. Liguori, D. Marinucci, S. Matarrese, L. Moscardini, and N. Vittorio. The integrated bispectrum as a test of cosmic microwave background non-Gaussianity : detection power and limits on  $f_{NL}$  with WMAP data. *Monthly Notices of the Royal Astronomical Society*, 369 :819–824, 2006.
- [60] A. Cabré, E. Gaztañaga, M. Manera, P. Fosalba, and F. Castander. Cross-correlation of Wilkinson Microwave Anisotropy Probe third-year data and the Sloan Digital Sky Survey DR4 galaxy survey : new evidence for dark energy. *Monthly Notices of the Royal Astronomical Society*, 372 :L23–L27, 2006.
- [61] S. Caillerie, M. Lachièze-Rey, J.-P. Luminet, R. Lehoucq, A. Riazuelo, and J. Weeks. A new analysis of the Poincaré dodecahedral space model. *Astronomy & Astrophysics*, 476 :691–696, 2007.
- [62] L. Cao, Y.-Q. Chu, and L.-Z. Fang. Cross-correlation between WMAP and 2MASS : non-Gaussianity induced by the SZ effect. *Monthly Notices of the Royal Astronomical Society*, 369 :645–654, 2006.
- [63] J.-F. Cardoso, M. Martin, J. Delabrouille, M. Betoule, and G. Patanchon. Component separation with flexible models. Application to the separation of astrophysical emissions. *preprint*, 2008.
- [64] L. Cayón, J. Jin, and A. Treaster. Higher Criticism statistic : detecting and identifying non-Gaussianity in the WMAP first-year data. *Monthly Notices of the Royal Astronomical Society*, 362 :826–832, 2005.
- [65] M. Cruz, E. Martinez-González, P. Vielva, and L. Cayon. Detection of a non-Gaussian spot in WMAP. *Monthly Notices of the Royal Astronomical Society*, 356 :29–40, 2005.
- [66] P. de Bernardis, P. Ade, J. Bock, J. Bond, J. Borrill, A. Boscaleri, K. Coble, B. Crill, G. De Gasperis, P. Farese, P. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. Hristov, A. Iacoangeli, A. Jaffe, A. Lange, L. Martinis, S. Masi, P. Mason, P. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, S. Prunet, S. Rao, G. Romeo, J. Ruhl, F. Scaramuzzi, D. Sforna, and N. Vittorio. A flat Universe from high-resolution maps of the cosmic microwave background radiation. *Nature*, 404 :955–959, 2000.
- [67] J. Delabrouille and J.-F. Cardoso. Diffuse source separation in CMB observations. *preprint*, 2007.
- [68] J. Delabrouille, J.-F. Cardoso, M. Le Jeune, M. Betoule, G. Faÿ, and F. Guilloux. A full sky, low foreground, high resolution CMB map from WMAP. *soumis*, 2008.

- [69] J. Delabrouille, J.-F. Cardoso, and G. Patanchon. Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis. *Monthly Notices of the Royal Astronomical Society*, 346 :1089–1102, 2003.
- [70] J.-F. Denisse, E. Le Roux, and J.-L. Steinberg. Observation du rayonnement galactique sur la longueur d’onde de 33cm. *Comptes Rendus de l’Académie des Sciences*, 240 :278–280, 1955.
- [71] R. Dicke, P. Peebles, P. Roll, and D. Wilkinson. Cosmic black-body radiation. *The Astrophysical Journal*, 142 :414–419, 1965.
- [72] S. Dodelson. *Modern Cosmology*. Academic Press, 2004.
- [73] A. Doroshkevich, P. Naselsky, O. Verkhodanov, D. Novikov, V. Turchaninov, I. Novikov, P. Christensen, and L. Chiang. Gauss–Legendre Sky Pixelization (GLESP) for CMB maps. *International Journal of Modern Physics D*, 14 :275, 2005.
- [74] G. Efstathiou. Myths and truths concerning estimation of power spectra : the case for a hybrid estimator. *Monthly Notices of the Royal Astronomical Society*, 349(2) :603–626, 2004.
- [75] H. Eriksen, A. Banday, K. Górski, and P. Lilje. On Foreground Removal from the Wilkinson Microwave Anisotropy Probe Data by an Internal Linear Combination Method : Limitations and Implications. *The Astrophysical Journal*, 612 :633–646, 2004.
- [76] H. Eriksen, C. Dickinson, J. Jewell, A. Banday, K. Gorski, and C. Lawrence. The joint large-scale foreground-CMB posteriors of the 3-year WMAP data. *preprint*, 2007.
- [77] H Eriksen, I. O’Dwyer, J. Jewell, B. Wandelt, D. Larson, K. Górski, S. Levin, A. Banday, and P. Lilje. Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling. *The Astrophysical Journal Supplement Series*, 155 :227–241, 2004.
- [78] G. Faÿ, F. Guilloux, M. Betoule, J.-F. Cardoso, J. Delabrouille, and M. Le Jeune. CMB power spectrum estimation using wavelets. *soumis*, 2008.
- [79] D. Finkbeiner. A Full-Sky  $H\alpha$  Template for Microwave Foreground Prediction. *The Astrophysical Journal Supplement Series*, 146 :407–415, 2003.
- [80] D. Finkbeiner, M. Davis, and D. Schlegel. Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS. *The Astrophysical Journal*, 524 :867–886, 1999.
- [81] P. Fosalba and E. Gaztañaga. Measurement of the gravitational potential evolution from the cross-correlation between WMAP and the APM Galaxy Survey. *Monthly Notices of the Royal Astronomical Society*, 350 :L37–L41, 2004.
- [82] B. Gold, C. Bennett, R. Hill, G. Hinshaw, N. Odegard, L. Page, D. Spergel, J. Weiland, J. Dunkley, M. Halpern, N. Jarosik, A. Kogut, E. Komatsu, D. Larson, S. Meyer, M. Nolta, E. Wollack, and E. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Galactic Foreground Emission. *preprint*, 2008.
- [83] J. González-Nuevo, F. Argüeso, M. López-Caniego, L. Toffolatti, J. Sanz, P. Vielva, and D. Herranz. The Mexican hat wavelet family : application to point-source detection in cosmic microwave background maps. *Monthly Notices of the Royal Astronomical Society*, 369 :1603–1610, 2006.

- [84] K. Górski, E. Hivon, A. Banday, B. Wandelt, F. Hansen, M. Reinecke, and M. Bartelmann. HEALPix : A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *The Astrophysical Journal*, 622 :759–771, 2005. Package available at <http://healpix.jpl.nasa.gov>.
- [85] P. Henry. Isotropy of the 3 K Background. *Nature*, 231 :516–518, 1971.
- [86] G. Hinshaw, M. Nolta, C. Bennett, R. Bean, O. Dore, M. Greason, M. Halpern, R. Hill, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, N. Odegard, S. Meyer, L. Page, H. Peiris, D. Spergel, G. Tucker, L. Verde, J. Weiland, E. Wollack, and E. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Temperature Analysis. *The Astrophysical Journal Supplement Series*, 170 :288–334, 2007.
- [87] G. Hinshaw, D. N. Spergel, L. Verde, R. S. Hill, S. S. Meyer, C. Barnes, C. L. Bennett, M. Halpern, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : The Angular Power Spectrum. *The Astrophysical Journal Supplement Series*, 148 :135–159, 2003.
- [88] G. Hinshaw, J. Weiland, R. Hill, N. Odegard, D. Larson, C. Bennett, J. Dunkley, B. Gold, M. Greason, N. Jarosik, E. Komatsu, M. Nolta, L. Page, D. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. Meyer, G. Tucker, and E. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Data Processing, Sky Maps, and Basic Results. *preprint*, 2008.
- [89] E. Hivon, K. Górski, C. Netterfield, B. Crill, S. Prunet, and F. Hansen. MASTER of the cosmic microwave background anisotropy power spectrum : A fast method for statistical analysis of large and complex cosmic microwave background data sets. *The Astrophysical Journal*, 567 :2–17, 2002.
- [90] W. Hu and S. Dodelson. Cosmic microwave background anisotropies. *Annual Review of Astronomy and Astrophysics*, 40 :171–216, 2002.
- [91] K. Huffenberger, H. Eriksen, F. Hansen, A. Banday, and K. Gorski. The scalar perturbation spectral index  $n_s$  : WMAP sensitivity to unresolved point sources. *preprint*, 2007.
- [92] J. Jewell, H. Eriksen, B. Wandelt, I. O’Dwyer, G. Huey, and K. Gorski. A Markov Chain Monte Carlo Algorithm for analysis of low signal-to-noise CMB data. *preprint*, 2008.
- [93] W. Jones, P. Ade, J. Bock, J. Bond, J. Borrill, A. Boscaleri, P. Cabella, C. Contaldi, B. Crill, P. de Bernardis, G. De Gasperis, A. de Oliveira-Costa, G. De Troia, G. di Stefano, E. Hivon, A. Jaffe, T. Kisner, A. Lange, C. MacTavish, S. Masi, P. Matuszewska, A. Melchiorri, T. Montroy, P. Natoli, C. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, G. Polenta, S. Prunet, S. Ricciardi, G. Romeo, J. Ruhl, P. Santini, M. Tegmark, M. Veneziani, and N. Vittorio. A Measurement of the Angular Power Spectrum of the CMB Temperature Anisotropy from the 2003 Flight of BOOMERANG. *The Astrophysical Journal*, 647 :823–832, 2006.
- [94] J. Kim, P. Naselsky, and P. Christensen. CMB map derived from the WMAP data through Harmonic Internal Linear Combination. *preprint*, 2008.

- [95] A. Klypin, M. Sazhin, I. Strukov, and D. Skulachev. Limits on Microwave Background Anisotropies - the Relikt Experiment. *Soviet Astronomy Letters*, 13 :104–+, 1987.
- [96] A. Kosowsky. The Atacama Cosmology Telescope project : A progress report. *New Astronomy Review*, 50 :969–976, 2006.
- [97] A. Kozowsky. The cosmic microwave background. In S. Bonometto, V. Gorini, and U. Moschella, editors, *Modern Cosmology*, pages 219–263. Institute of Physics Publishing, 2002.
- [98] M. Lachièze-Rey. *Initiation à la cosmologie (Cours de Master)*. Dunod, 4e edition, 2004.
- [99] J.-M. Lamarre, J.-L. Puget, F. Bouchet, P. Ade, A. Benoit, J. Bernard, J. Bock, P. de Bernardis, J. Charra, F. Couchot, J. Delabrouille, G. Efstathiou, M. Giard, G. Guyot, A. Lange, B. Maffei, A. Murphy, F. Pajot, M. Piat, I. Ristorcelli, D. Santos, R. Sudiwala, J.-F. Sygnet, J.-P. Torre, V. Yurchenko, and D. Yvon. The Planck High Frequency Instrument, a third generation CMB experiment, and a full sky submillimeter survey. *New Astronomy Review*, 47 :1017–1024, 2003.
- [100] S. Leach, J.-F. Cardoso, C. Baccigalupi, R. Barreiro, M. Betoule, J. Bobin, A. Bonaldi, G. de Zotti, J. Delabrouille, C. Dickinson, H. Eriksen, J. González-Nuevo, F. Hansen, D. Herranz, M. Le Jeune, M. López-Caniego, E. Martínez-González, M. Massardi, J. B. Melin, G. Miville-Deschênes, M. Patanchon, S. Prunet, S. Ricciardi, E. Salerno, J. Sanz, J.-L. Starck, F. Stivoli, V. Stolyarov, R. Stompor, and P. Vielva. Component separation methods for the Planck mission. *preprint*, 2008.
- [101] A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *The Astrophysical Journal*, 538 :473–476, 2000.
- [102] P. Lubin, T. Villela, G. Epstein, and G. Smoot. A map of the cosmic background radiation at 3 millimeters. *The Astrophysical Journal Letters*, 298 :L1–L5, 1985.
- [103] D. Maino, S. Donzelli, A. Banday, F. Stivoli, and C. Baccigalupi. Cosmic microwave background signal in Wilkinson Microwave Anisotropy Probe three-year data with FASTICA. *Monthly Notices of the Royal Astronomical Society*, 374 :1207–1215, 2007.
- [104] D. Maino, A. Farusi, C. Baccigalupi, F. Perrotta, A. J. Banday, L. Bedini, C. Burigana, G. De Zotti, K. M. Górski, and E. Salerno. All-sky astrophysical component separation with Fast Independent Component Analysis (FASTICA). *Monthly Notices of the Royal Astronomical Society*, 334 :53–68, 2002.
- [105] D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella, G. Kerkycharian, P. Natoli, D. Picard, and N. Vittorio. Spherical Needlets for CMB Data Analysis. *Monthly Notices of the Royal Astronomical Society*, 383(2) :539–545, 2008.
- [106] S. Masi, P. Ade, J. Bock, J. Bond, J. Borrill, A. Boscaleri, P. Cabella, C. Contaldi, B. Crill, P. de Bernardis, G. de Gasperis, A. de Oliveira-Costa, G. de Troia, G. di Stefano, P. Ehlers, E. Hivon, V. Hristov, A. Iacoangeli, A. Jaffe, W. Jones, T. Kisner, A. Lange, C. MacTavish, C. Marini Bettolo, P. Mason, P. Mausekopf, T. Montroy, F. Nati, L. Nati, P. Natoli, C. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, G. Polenta, S. Prunet, S. Ricciardi, G. Romeo, J. Ruhl, P. Santini, M. Tegmark, E. Torbet, M. Veneziani, and N. Vittorio. Instrument, method, brightness, and polarization maps from the 2003 flight of BOOMERanG. *Astronomy & Astrophysics*, 458 :687–716, 2006.

- [107] J. McEwen, M. Hobson, A. Lasenby, and D. Mortlock. A high-significance detection of non-Gaussianity in the WMAP 3-yr data using directional spherical wavelets. *Monthly Notices of the Royal Astronomical Society*, 371 :L50–L54, 2006.
- [108] J. McEwen, M. Hobson, A. Lasenby, and D. Mortlock. Non-Gaussianity detections in the Bianchi  $VII_h$  corrected WMAP one-year data made with directional spherical wavelets. *Monthly Notices of the Royal Astronomical Society*, 369 :1858–1868, 2006.
- [109] J. McEwen, Y. Wiaux, M. Hobson, P. Vandergheynst, and A. Lasenby. Probing dark energy with steerable wavelets through correlation of wmap and nvss local morphological measures. *Monthly Notices of the Royal Astronomical Society*, 384 :1289–1300, 2008.
- [110] C. Miller, R. Nichol, C. Genovese, and L. Wasserman. A Nonparametric Analysis of the Cosmic Microwave Background Power Spectrum. *The Astrophysical Journal*, 565 :L67–L70, 2002.
- [111] M. Miville-Deschênes and G. Lagache. IRIS : A New Generation of IRAS Maps. *The Astrophysical Journal Supplement Series*, 157 :302–323, 2005.
- [112] P. Mukherjee and Y. Wang. Wavelets and Wilkinson Microwave Anisotropy Probe Non-Gaussianity. *The Astrophysical Journal*, 613 :51–60, September 2004.
- [113] A. Niarchou and A. Jaffe. Imprints of Spherical Nontrivial Topologies on the Cosmic Microwave Background. *Physical Review Letters*, 99(8) :081302–+, August 2007.
- [114] M. Nolta, J. Dunkley, R. Hill, G. Hinshaw, E. Komatsu, D. Larson, L. Page, D. Spergel, C. Bennett, B. Gold, N. Jarosik, N. Odegard, J. Weiland, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. Meyer, G. Tucker, and E. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Angular Power Spectra. *preprint*, 2008.
- [115] M. Nolta, E. Wright, L. Page, C. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. Meyer, D. Spergel, G. Tucker, and E. Wollack. First Year Wilkinson Microwave Anisotropy Probe Observations : Dark Energy Induced Correlation with Radio Sources. *The Astrophysical Journal*, 608 :10–15, 2004.
- [116] P. Oxley, P. Ade, C. Baccigalupi, P. deBernardis, H.-M. Cho, M. Devlin, S. Hanany, B. Johnson, T. Jones, A. Lee, T. Matsumura, A. Miller, M. Milligan, T. Renbarger, H. Spieler, R. Stompor, G. Tucker, and M. Zaldarriaga. The EBEX experiment. In M. Strojnik, editor, *Infrared Spaceborne Remote Sensing XII.*, volume 5543 of *Proceedings of the SPIE*, pages 320–331, 2004.
- [117] L. Page, C. Barnes, G. Hinshaw, D. Spergel, J. Weiland, E. Wollack, C. Bennett, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. Meyer, G. Tucker, and E. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Beam Profiles and Window Functions. *The Astrophysical Journal Supplement Series*, 148 :39–50, 2003.
- [118] C. Park, C. Park, and J. Gott. Cleaned 3 Year Wilkinson Microwave Anisotropy Probe Cosmic Microwave Background Map : Magnitude of the Quadrupole and Alignment of Large-Scale Modes. *The Astrophysical Journal*, 660 :959–975, 2007.
- [119] P. Peebles. Statistical analysis of catalogs of extragalactic objects. I. Theory. *The Astrophysical Journal*, 185 :413–440, 1973.

- [120] A. Penzias and R. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal*, 142 :419–421, 1965.
- [121] D. Pietrobon, A. Balbi, and D. Marinucci. Integrated Sachs-Wolfe effect from the cross correlation of WMAP 3year and the NRAO VLA sky survey data : New results and constraints on dark energy. *Physical Review D*, 74(4) :043524–+, 2006.
- [122] G. Polenta, D. Marinucci, A. Balbi, P. de Bernardis, E. Hivon, S. Masi, P. Natoli, and N. Vittorio. Unbiased estimation of an angular power spectrum. *Journal of Cosmology and Astroparticle Physics*, 11 :1–+, 2005.
- [123] A. Rassat, K. Land, O. Lahav, and F. Abdalla. Cross-correlation of 2MASS and WMAP 3 : Implications for the integrated Sachs-Wolfe effect. *Monthly Notices of the Royal Astronomical Society*, 377 :1085–1094, 2007.
- [124] C. Reichardt, P. Ade, J. Bock, J. Bond, J. Brevik, C. Contaldi, M. Daub, J. Dempsey, J. Goldstein, W. Holzapfel, C. Kuo, A. Lange, M. Lueker, M. Newcomb, J. Peterson, J. Ruhl, M. Runyan, and Z. Staniszewski. High resolution CMB power spectrum from the complete ACBAR data set. *The Astrophysical Journal*, 2008. (In press).
- [125] M. Runyan, P. Ade, R. Bhatia, J. Bock, M. Daub, J. Goldstein, C. Haynes, W. Holzapfel, C. Kuo, A. Lange, J. Leong, M. Lueker, M. Newcomb, J. Peterson, C. Reichardt, J. Ruhl, G. Sirbi, E. Torbet, C. Tucker, A. Turner, and D. Woolsey. ACBAR : The Arcminute Cosmology Bolometer Array Receiver. *The Astrophysical Journal Supplement Series*, 149 :265–287, 2003.
- [126] M. Runyan, P. Ade, J. Bock, J. Bond, C. Cantalupo, C. Contaldi, M. Daub, J. Goldstein, P. Gomez, W. Holzapfel, C. Kuo, A. Lange, M. Lueker, M. Newcomb, J. Peterson, D. Pogosyan, A. Romer, J. Ruhl, E. Torbet, and D. Woolsey. First results from the arcminute cosmology bolometer array receiver. *New Astronomy Review*, 47 :915–923, 2003.
- [127] R. Sachs and A. Wolfe. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *The Astrophysical Journal*, 147 :73–90, 1967.
- [128] R. Saha, S. Prunet, P. Jain, and T. Souradeep. CMB anisotropy power spectrum using linear combinations of WMAP maps. *Physical Review D*, 78(2) :023003–+, 2008.
- [129] D. Samtleben and for the QUIET collaboration. QUIET - Measuring the CMB polarization with coherent detector arrays. *preprint*, 2008.
- [130] U. Seljak and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *The Astrophysical Journal*, 469 :437–+, 1996.
- [131] J. Silk. Fluctuations in the Primordial Fireball. *Nature*, 215 :1155–+, 1967.
- [132] G. Smoot, C. Bennett, A. Kogut, E. Wright, J. Aymon, N. Boggess, E. Cheng, G. de Amici, S. Gulkis, M. Hauser, G. Hinshaw, P. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. Meyer, S. Moseley, T. Murdock, L. Rokke, R. Silverberg, L. Tenorio, R. Weiss, and D. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. *The Astrophysical Journal*, 396 :L1–L5, 1992.

- [133] J.-L. Starck, N. Aghanim, and O. Forni. Detection and discrimination of cosmological non-gaussian signatures by multi-scale methods. *Astronomy & Astrophysics*, 416 :9, 2004.
- [134] J.-L. Starck, Y. Moudden, P. Abrial, and M. Nguyen. Wavelets, ridgelets and curvelets on the sphere. *Astronomy & Astrophysics*, 446 :1191–1204, 2006.
- [135] I. Strukov and D. Skulachev. Deep-Space Measurements of the Microwave Background Anisotropy - First Results of the Relikt Experiment. *Soviet Astronomy Letters*, 10 :1–+, 1984.
- [136] R. Sunyaev and Y. Zeldovich. Small-Scale Fluctuations of Relic Radiation. *Astrophysics and Space Science*, 7 :3–19, 1970.
- [137] M. Tegmark. How to measure cmb power spectra without losing information. *Physical Review D*, 55(10) :5895–5907, 1997.
- [138] M. Tegmark. Removing Real-World Foregrounds from Cosmic Microwave Background Maps. *The Astrophysical Journal*, 502 :1–+, 1998.
- [139] M. Tegmark, A. de Oliveira-Costa, and A. Hamilton. High resolution foreground cleaned CMB map from WMAP. *Physical Review D*, 68(12) :123523–+, 2003.
- [140] R. Tojeiro, P. Castro, A. Heavens, and S. Gupta. Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe data using the peak-peak correlation function. *Monthly Notices of the Royal Astronomical Society*, 365 :265–275, 2006.
- [141] L. Valenziano, M. Sandri, G. Morgante, C. Burigana, M. Bersanelli, R. Butler, F. Cuttaia, F. Finelli, E. Franceschi, M. Galaverni, A. Gruppuso, M. Malaspina, N. Mandolesi, A. Mennella, F. Paci, L. Popa, P. Procopio, L. Stringhetti, L. Terenzi, M. Tomasi, F. Villa, and J. Zuccarelli. The low frequency instrument on-board the Planck satellite : Characteristics and performance. *New Astronomy Review*, 51 :287–297, 2007.
- [142] P. Vielva, E. Martinez-Gonzalez, R. Barreiro, J. Sanz, and L. Cayon. Detection of non-Gaussianity in the WMAP 1-year data using spherical wavelets. *The Astrophysical Journal*, 609 :22, 2004.
- [143] R. Vio and P. Andreani. A Statistical Analysis of the "Internal Linear Combination" Method in Problems of Signal Separation as in CMB Observations. *preprint*, 2008.
- [144] B. Wandelt, D. Larson, and A. Lakshminarayanan. Global, exact cosmic microwave background data analysis using gibbs sampling. *Physical Review D*, 70(8) :083511, 2004.
- [145] Y. Wiaux, J. McEwen, P. Vandergheynst, and O. Blanc. Exact reconstruction with directional wavelets on the sphere. *Monthly Notices of the Royal Astronomical Society*, 388 :770–788, 2008.
- [146] E. Wright, X. Chen, N. Odegard, C. Bennett, R. Hill, G. Hinshaw, N. Jarosik, E. Komatsu, M. Nolte, L. Page, D. Spergel, J. Weiland, E. Wollack, J. Dunkley, B. Gold, M. Halpern, A. Kogut, D. Larson, M. Limon, S. Meyer, and G. Tucker. The Wilkinson Microwave Anisotropy Probe (WMAP) Source Catalog. *preprint*, 2008.
- [147] A. Yadav and B. Wandelt. Detection of primordial non-Gaussianity (fNL) in the WMAP 3-year data at above 99.5% confidence. *Physical Review Letters*, 100(18) :181301–+, 2008.

## Statistiques et Probabilités

- [148] Z. Bai, C. Rao, and L. Zhao. Kernel estimators of density function of directional data. *Journal of Multivariate Analysis*, 27(1) :24–39, 1988.
- [149] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard. Subsampling Needlet Coefficients on the Sphere. *preprint*, 2007.
- [150] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard. Adaptive density estimation for directional data using needlets. *preprint*, 2008.
- [151] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard. Asymptotics for spherical needlets. *Annals of Statistics*, 2008. (In press).
- [152] P. Baldi, G. Kerkycharian, D. Marinucci, and D. Picard. High frequency asymptotics for wavelet-based tests for Gaussianity and isotropy on the torus. *Journal of Multivariate Analysis*, 99 :606–636, 2008.
- [153] P. Baldi and D. Marinucci. Some characterizations of the spherical harmonics coefficients for isotropic random fields. *Statistics and Probability Letters*, 77 :490–496, 2007.
- [154] J.-M. Bardet and P. Bertrand. Wavelet analysis of a continuous-time gaussian process observed at random times and its application to the estimation of the spectral density. *preprint*, 2008.
- [155] P. Cabella and D. Marinucci. Statistical challenges in the analysis of cosmic microwave background radiation. *Annals of Applied Statistics*, 2008. (In press).
- [156] A. Cohen and R. Jones. Regression on a random field. *Journal of the American Statistical Association*, 64 :1172–1182, 1969.
- [157] G. Faÿ and F. Guillaoux. Consistency of a needlet spectral estimator on the sphere. *soumis*, 2008.
- [158] N. Fisher, T. Lewis, and B. Embleton. *Statistical analysis of spherical data*. Cambridge University Press, 1987.
- [159] C. Genovese, C. Miller, R. Nichol, M. Arjunwadkar, and L. Wasserman. Nonparametric Inference for the Cosmic Microwave Background. *Statistical Science*, 19 :308–321, 2004.
- [160] P. Hall, G. Watson, and J. Cabrera. Kernel density estimation with spherical data. *Biometrika*, 74(4) :751–762, 1987.
- [161] D. Healy, H. Hendriks, and P. Kim. Spherical deconvolution. *Journal of Multivariate Analysis*, 67 :1–22, 1998.
- [162] H. Hendriks. Application of fast spherical Fourier transform to density estimation. *Journal of Multivariate Analysis*, 84(2) :209–221, 2003.
- [163] R. Jones. Stochastic processes on a sphere. *Annals of Mathematical Statistics*, 34 :213–218, 1963.
- [164] M. Jun and M. Stein. An approach to producing space-time covariance functions on spheres. *Technometrics*, 49(4) :468–479, 2007.



- [165] M. Jun and M. Stein. Nonstationary covariance models for global data. *Annals of Applied Statistics*, 2008. (In press).
- [166] P. Kim and J.-Y. Koo. Optimal spherical deconvolution. *Journal of Multivariate Analysis*, 80 :21–42, 2002.
- [167] J. Klemelä. Asymptotic minimax risk for the white noise model on the sphere. *Scandinavian Journal of Statistics*, 26(3) :465–473, 1999.
- [168] J. Klemelä. Lower bounds for the asymptotic minimax risk with spherical data. *Journal of Statistical Planning and Inference*, 113(1) :113–136, 2003.
- [169] J.-Y. Koo and P. Kim. Sharp adaptation for spherical inverse problems with applications to medical imaging. *Journal of Multivariate Analysis*, 99 :165–190, 2008.
- [170] X. Lan and D. Marinucci. On the dependence structure of wavelet coefficients for spherical random fields. *preprint*, 2008.
- [171] X. Lan and D. Marinucci. The needlet bispectrum. *Electronic Journal of Statistics*, 2 :332–367, 2008.
- [172] K. Mardia. *Statistics of directional data*. Probability and Mathematical Statistics, No. 13. Academic Press, 1972.
- [173] K. Mardia and P. Jupp. *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2000.
- [174] D. Marinucci. Testing for non-Gaussianity on cosmic microwave background radiation : a review. *Statistical Science*, 19(2) :294–307, 2004.
- [175] A. Mayeli. Asymptotic uncorrelation for mexican needlets. *preprint*, 2008.
- [176] F. Roueff and M. Taqqu. Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *preprint*, 2008.
- [177] R. Roy. Estimation of the covariance function of a homogeneous process on the sphere. *Annals of Statistics*, 1(4) :780–785, 1973.
- [178] R. Roy. Spectral analysis for a random process on the sphere. *Annals of the Institute of Statistical Mathematics*, 28(1) :91–97, 1974.
- [179] F. Ruymgaart. Strong uniform convergence of density estimators on spheres. *Journal of Statistical Planning and Inference*, 23(1) :45–52, 1989.
- [180] G. Upton and B. Fingleton. *Spatial Data Analysis by Example. Vol 2 : Categorical and Directional Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, 1989.
- [181] A. Yaglom. Second-order homogeneous random fields. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, pages 593–622, 1961.

## Divers, Autres applications

- [182] F. Dahlen and F. Simons. Spectral estimation on a sphere in geophysics and cosmology. *Geophysical Journal International*, 2008. (In press).
- [183] W. Freeden and M. Schreiner. An integrated wavelet concept of physical geodesy. *Journal of Geodesy*, 72(5) :259–281, 1998.
- [184] M. Holschneider, A. Chambodut, and M. Manda. From global to regional analysis of magnetic field on the sphere using wavelet frames. *Physics of the Earth and Planetary Interiors*, 135 :107–124, 2003.
- [185] H.-S. Oh. Spherical wavelets and their application to meteorological data. In O. Leuangthong and C. Deutsch, editors, *Geostatistics Banff 2004*, Quantitative Geology and Geostatistics, pages 819–831. Springer, 2006.
- [186] H.-S. Oh and T.-H. Li. Estimation of global temperature fields from scattered observations by a spherical-wavelet-based spatially adaptive method. *Journal of the Royal Statistical Society B*, 67 :1–22, 1998.
- [187] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty — V : The discrete case. *Bell System Technical Journal*, 57 :1371–1429, 1978.
- [188] D. Slepian and H. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty — I. *Bell System Technical Journal*, 40(1) :43–63, 1960.
- [189] E. Thébault. Global lithospheric magnetic field modelling by successive regional analysis. *Earth, Planets and Space*, 58 :485–495, 2006.
- [190] M. Wieczorek and F. Simons. Localized spectral analysis on the sphere. *Geophysical Journal International*, 162(3) :655–675, 2005.
- [191] P. Yu, P. Grant, Y. Qi, X. Han, F. Segonne, R. Pienaar, E. Busa, J. Pacheco, N. Makris, R. Buckner, P. Golland, and B. Fischl. Cortical surface shape analysis based on spherical wavelets. *IEEE Transactions on Medical Imaging*, 26(4) :582–597, 2007.



**Résumé :** Cette thèse porte sur l'utilisation d'ondelettes de seconde génération pour l'étude statistique de champs aléatoires sphériques. Parmi les phénomènes modélisables par un tel champ, on s'intéressera en particulier au Fond diffus cosmologique (CMB).

La localisation des needlets (récente construction d'ondelettes) sur la sphère est étudiée et optimisée en terme de concentration spatiale et d'estimation statistique. Ces fonctions sont ensuite utilisées pour construire un nouvel estimateur du spectre de puissance angulaire. L'examen des propriétés de cet estimateur, d'un point de vue théorique (dans l'asymptotique des hautes fréquences angulaires) et pratique, montre qu'il améliore les méthodes existantes dans un modèle réaliste comportant des données manquantes et un bruit hétéroscédastique. À côté de l'estimation spectrale, l'utilisation des needlets est également introduite dans un problème de séparation de sources.

Après quatre chapitres introductifs (dédiés respectivement aux aspects physiques, analytiques et statistiques de l'étude du CMB, puis à une présentation d'ensemble des résultats), quatre articles de revue (en collaboration) sont présentés : *Practical wavelet design on the sphere* ; *CMB power spectrum estimation using wavelets* ; *Spectral estimation on the sphere with needlets: high frequency asymptotics* et *A full sky, low foreground, high resolution CMB map from WMAP*.

**Mots-clés :** Fond diffus cosmologique (Rayonnement fossile) – Analyse sur la sphère – Ondelettes (needlets) – Champ (Processus) stationnaire (isotrope) – Spectre de puissance angulaire – Estimation spectrale – Données manquantes – Modèle hétéroscédastique – Asymptotique des hautes fréquences – Séparation de sources.

---

**Abstract:** This thesis focuses on the use of second generation wavelet frames for the statistical study of spherical random fields. Among phenomena modeled by such a field, we are specifically interested in the Cosmic Microwave Background (CMB).

The localization of needlets (recent construction of wavelets) on the sphere is considered and optimized in terms of spatial concentration and statistical estimation. These functions are then used to define a new estimator of the angular power spectrum. The study of this estimator's theoretical (in the high-frequency asymptotics) and practical properties shows that it improves existing methods in a realistic model including missing data and heteroscedastic noise. Beside spectral estimation, the use of needlets is also introduced in a component separation framework.

Four introductory chapters are devoted to physical, analytical and statistical aspects of the CMB, and to the presentation of our results. Then, we present four articles (in collaboration) : *Practical wavelet design on the sphere* ; *CMB power spectrum estimation using wavelets* ; *Spectral estimation on the sphere with needlets: high frequency asymptotics* and *A full sky, low foreground, high resolution CMB map from WMAP*.

**Key words:** Cosmic Microwave Background (CMB) – Analysis on the Sphere – Wavelets (needlets) – Stationary (Isotropic) field (process) – Angular power spectrum – Spectral estimation – Missing data – Heteroscedastic model – High-frequency asymptotics – Component separation.